

Master 2 professionnel Bioingénierie options Biotechnologies végétales  
Université Paul Sabatier. Toulouse III

*Cartographie génétique de gènes candidats pour  
la qualité du bois chez l'Eucalyptus.*



*Tristan COSTECALDE*

Responsable de stage :  
**Jean-Marc GION**



Laboratoire de Génétique  
Programme Arbres et Plantations  
CIRAD

Soutenu à Toulouse le  
13 septembre 2007

## SOMMAIRE

<b>1) INTRODUCTION</b>	<b>1</b>
<b>2) MATERIEL et METHODES</b>	<b>7</b>
2.1 - Matériel végétal	7
2.2 - Caractéristiques des EST étudiées	8
2.3 - Amplification des EST par PCR dirigée	9
2.4 - Mise en évidence du polymorphisme	9
2.4.1 - Séquençage	10
2.4.2 - Révélation du polymorphisme par PCR en temps réel	10
2.4.3 - Révélation du polymorphisme par la technique SSCP	12
2.4.3.1 - La SSCP sur LI-COR	12
2.4.3.2 - La SSCP sur séquenceur ABI	14
2.5 - Cartographie génétique des EST polymorphes	15
<b>3) RESULTATS</b>	<b>17</b>
3.1 - Amplification des EST chez les géniteurs 9.21 et 14.144	17
3.2 - Mise en évidence de la variabilité chez les parents	17
3.3 - Mise en évidence du polymorphisme chez les descendants	19
3.3.1 - La technique PCR en temps réel pour le génotypage haut débit	19
3.3.2 - Génotypage par SSCP sur LI-COR et SSCP sur ABI	20
3.3.2.1 - Premiers tests effectués pour la SSCP sur LI-COR et ABI	20
3.3.2.2 - Sélection des EST polymorphes par SSCP sur LI-COR	21
3.3.2.3 - Génotypage des EST polymorphes sur la descendance <i>E. urophylla</i> x <i>E. grandis</i>	22
3.4 - Localisation des EST polymorphes sur les cartes génétiques	23
<b>4) DISCUSSION</b>	<b>24</b>
-La conservation des séquences entre espèces	24
-La PCR en temps réel	24
-La SSCP sur LI-COR et ABI	26
-La cartographie d'EST	28
<b>5) CONCLUSION ET PERSPECTIVES</b>	<b>31</b>

## BIBLIOGRAPHIE

## ANNEXES



## REMERCIEMENTS :

Je tiens à exprimer toute ma gratitude et ma reconnaissance :

A Jean - Marc Gion, mon maître de stage, pour son encadrement robuste au sein de la Dream Team « Eucalpytus ». Vraiment Merci, pour cet enseignement dont « rigueur » et « perfectionnisme » sont les maîtres mots.

A Christophe Plomion pour sa disponibilité a toute épreuve.

A François Hubert mon ami et collègue de bureau, pour sa patience et son dévouement. Vraiment un grand merci. Grâce à toi, Microsoft Word n'aura plus de secret pour moi.

A Eric pour son soutien en manips, sa gentillesse et son humour massacrant. Sacré Eric, les soirées chez toi resteront inoubliables ! Garde ton jeu de jambes papa et surtout ne change rien !

Au plus beau couple de l'année : Séverine et Briçou, pour m'avoir réellement soutenu dans les moments difficiles et pour tous ces bons moments de complicité passés ensemble.

A Olivier pour m'avoir fait découvrir les bons plans de Bordeaux.

A Delphine et Julie pour les fabuleuses soirées passées à l'Azuli.

A M<sup>R</sup> « C'aumeil » pour sa gentillesse et son investissement dans les manips de PCR en temps réel. Vraiment merci !

A Emilie, pour m'avoir accompagné dans les manips SSCP.

A Frank Salin, Christophe Hubert, Jorge, Erwan et Jerome pour leur aide en laboratoire.

A Maïmiti pour sa bonne humeur permanente.

A Florian et Klem pour leur soutien moral et les soirées pizza.

A mes parents, merci.

## Abréviations

- AFLP	Amplified fragment-length polymorphism
- APS	Persulfate d'ammonium
- ATMAB	Alkyltrimethylammonium bromides
- ADN	Acide désoxyribonucléique
- ADNc	Acide désoxyribonucléique complémentaire
- BET	Bromure d'éthidium
- BLAST	Basic local alignment search tool
- CAPS	Cleaved amplified polymorphic sequence
- CIRAD	Centre de Coopération Internationale en Recherche Agronomique pour le Développement
- cM	Centimorgan
- db	double brin
- DMSO	Dimethyl Sulfoxide
- dNTP	Désoxynucléotide triphosphate
- ddNTP	Didésoxyribonucléotide triphosphate
- EDTA	Acide éthylène diamine tétracétique
- EST	Expressed sequence tag
- FAO	Food and Agriculture organization
- GL	Groupe de liaison
- h	Heure
- HCl	Acide chlorhydrique
- INDEL	Insertion, deletion
- KCl	chlorure de potassium
- LOD	Log10 of the odds ratio
- M	Molaire
- MDE	Mutation Detection Enhancement
- MgCl <sup>2</sup>	Chlorure de magnesium
- min	Minute
- MIPS	Munich Information Centre for Protein Sequences
- ml	Millilitre
- mM	Millimolaire
- NaCl	sodium chloride
- ng	Nanogramme
- pb	Paire de bases
- RT PCR	Reverse Transcription PCR
- PCR	Polymerase chain reaction
- QTL	Quantitative trait loci
- RAPD	Random amplified polymorphic DNA
- RFLP	Restriction fragment length polymorphism
- SAM	Sélection assistée par marqueurs
- sec	Seconde
- SNP	Single nucleotide polymorphism
- SSCP	Single strand conformation polymorphism
- SSH	Suppressive subtractive hybridization
- SSR	Single sequence repeat ou microsatellite

- STS	Sequence-tagged site
- SYBR	Synergy Brands
- Taq	<i>Thermus aquaticus</i>
- TBE	Tris borate EDTA
- TEMED	Tetramethylethylenediamine
- T <sub>m</sub>	Melting temperature
- Tris HCL	Tris Hydrochloride
- UFR	Unité relative de fluorescence
- UV	Ultra violet
- V	Volt
- µg	Microgramme
- µl	Microlitre
- µmol	micromole



## 1) INTRODUCTION

Les eucalyptus sont natifs d'Australie. Le genre *Eucalyptus* appartient à la famille des *Myrtacées*. Il compte à lui seul plus de 700 espèces parmi lesquelles 13 sous-genres se distinguent. Le sous genre *Symphyomyrtus* est le plus représenté et se réparti sur tout le continent Australien, en Tasmanie, en Nouvelle-Guinée et dans l'Archipel de la Sonde. Il comprend huit sections dont trois contiennent l'essentiel des espèces cultivées : la section *Transversaria* (*E. grandis*, *E. saligna*, *E. urophylla*...), la section *Exsertaria* (*E. tereticornis*) et la section *Maidenaria* (*E. globulus*, *E. gunnii*). A l'heure actuelle, environ 80% de la production est utilisée comme bois d'énergie (bois de chauffage et charbon de bois) (Vigneron et Bouvet, 1997). Sur les 20% restants, la majeure partie est utilisée comme bois de trituration (pâte à papier et panneau de particules). Les caractéristiques de forme et la croissance rapide de l'eucalyptus, sont utilisées pour la production de bois de service (poteaux, perches, éléments de charpente en bois rond). En Australie, aire d'origine du genre, l'eucalyptus est aussi utilisé comme bois d'œuvre, mais cette utilisation reste rare dans les pays où il a été introduit. D'après la FAO, il y aurait actuellement 19 millions d'hectares plantés dans plus de 37 pays (Nguyen, 2006). L'utilisation industrielle de cet arbre repose sur ses propriétés sylvicoles intéressantes en plantation : une forte croissance dans le jeune âge permettant une exploitation rapide du bois et une capacité à coloniser des sols pauvres impropres à l'agriculture. A titre de comparaison l'âge d'exploitation d'*E. urophylla* se situe vers la septième année alors qu'il faut attendre 20 ans chez les peupliers et 40 à 80 ans chez les pins.

Les premiers travaux d'amélioration génétique de l'eucalyptus au Congo ont débuté en 1955, avec l'introduction de plusieurs espèces. Ces premiers travaux ont permis de sélectionner certaines espèces intéressantes et ont conduit à la mise en place de 5000 hectares de plantations (Vigneron et Delwaulle, 1990). L'introduction de ces espèces a engendré l'apparition dans les années 70 d'hybrides naturels aux caractéristiques de croissance et d'adaptation supérieures aux espèces introduites (Chaperon, 1977). Dès lors, la maîtrise du bouturage herbacé à partir de rejets de souche (Martin et Quillet, 1974) a permis une multiplication à grande échelle de ces hybrides. La mise au point de la pollinisation contrôlée en 1977 (Maillard, 1978) a alors permis de tester différentes combinaisons hybrides. Les formules hybrides qui se sont révélées être les plus prometteuses sont *E. urophylla* x *E. grandis* et *E. urophylla* x *E. pellita*. Dans le croisement *E. urophylla* x *E. grandis* au Congo, ces deux espèces sont fortement

complémentaires : *E. urophylla* est bien adapté aux conditions congolaises tandis qu'*E. grandis* est mal adapté mais présente de fortes potentialités de croissance. Depuis 1989, une stratégie de sélection récurrente réciproque a été adoptée (Vigneron, 1991). Cette méthode est basée sur l'amélioration conjointe de deux groupes d'individus l'un par rapport à l'autre afin d'obtenir des hybrides inter-groupes recombinaut différents caractères présents séparément dans chacun des groupes.

L'efficacité de la sélection classique chez les espèces forestières est limitée par deux contraintes majeures. D'une part, la mauvaise corrélation juvénile-adulte pour la plupart des caractères ciblés entraînant une évaluation tardive des caractères (Bouvet, 1995). D'autre part, les programmes d'amélioration des arbres nécessitent des surfaces considérables sur des périodes très longues (7 ans chez l'eucalyptus). Chez les arbres forestiers, deux grands axes de recherches visant à limiter ces problèmes sont développés :

- la transformation génétique couplée à une multiplication en masse du matériel végétal par embryogénèse somatique (Harvengt *et al.*, 2000) afin d'obtenir des génotypes d'intérêt industriel.
- l'utilisation de la sélection assistée par marqueurs visant à identifier et mettre à profit la variabilité génétique naturelle des gènes impliqués dans des caractères d'intérêt.

Jusqu'à présent, les programmes d'amélioration des espèces forestières se sont focalisés sur des caractères relevant de la sylviculture (croissance) et de l'adaptabilité (la résistance aux maladies, aux ravageurs). Il n'en a pas été de même pour l'étude des caractères de qualité du bois, qui nécessitait des méthodes d'évaluation destructives et coûteuses. Or, les gains sylvicoles se traduisent souvent par une diminution de la qualité du bois. Il est donc apparu nécessaire d'intégrer des notions de qualité du bois dans les programmes d'amélioration avec le développement de nouvelles méthodes non destructives et permettant une évaluation sur un grand nombre d'individus. La qualité du bois inclut une grande variété de caractéristiques pouvant se classer dans cinq catégories : technologiques (densité du bois), rhéologiques (retrait au séchage), chimiques (taux de lignine, cellulose et composés extractibles...), anatomiques (angle de microfibrilles, morphologie des fibres), papetières (rendement en pâte, extractibilité des lignines, indice de lignification...) (Eyles, 2003). De plus, les caractères d'intérêt varient en fonction de l'utilisation du produit. Ainsi pour la production papetière une teneur en cellulose élevée et une faible teneur en lignine sera recherchée (Wimmer *et al.*, 2002). A l'inverse une teneur élevée en lignine sera souhaitée pour le bois d'énergie de par son pouvoir calorifique



plus élevé que la cellulose (Raymond, 2002). Cependant dans tous les cas, la qualité du bois dépend essentiellement de la morphologie des fibres et des propriétés physico-chimiques de leurs parois secondaires (Miranda et Pereira, 2002). L'étude approfondie des mécanismes impliqués dans la formation du bois apparaît donc nécessaire pour une bonne maîtrise de la qualité du bois.

Dans cette optique, de nombreux travaux de recherche sur la xylogénèse ont été réalisés (Plomion *et al.*, 2001). Les travaux de génomique en foresterie ont commencé avec les projets de séquençages d'EST (Expressed Sequence Tag) à partir de xylème secondaire de pin et de peuplier (Allona *et al.*, 1998 ; Sterky *et al.*, 1998). L'intérêt des EST est d'isoler et d'identifier des gènes exprimés dans un tissu particulier, à un moment et dans un environnement précis. L'EST correspond à une séquence partielle d'ADNc (ARN reverse transcrit en ADN) de quelques centaines de nucléotides et représente ainsi la partie codante d'un gène (séquence sans intron). De grandes banques d'EST ont été développées chez les arbres. Les EST étudiées vont dépendre du caractère ciblé. Ainsi pour étudier des gènes impliqués dans la formation du bois, les transcrits seront généralement extraits dans le xylème (Allona *et al.*, 1998) ou dans le cambium (Bossinger *et al.*, 2000). Une fois isolées, ces EST sont identifiées par comparaison à des banques de données de séquences tel que GeneBank.

Les premiers travaux d'étude de gènes impliqués dans la qualité du bois étaient basés sur le prélèvement de xylème en différenciation et l'extraction de transcrits. Ces travaux ont permis à Allona *et al.* (1998) chez le pin, Sterky *et al.* (1998) chez le peuplier, d'isoler un grand nombre d'EST exprimées préférentiellement dans le xylème. Depuis, le séquençage d'EST se multiplie chez les espèces forestières notamment chez l'eucalyptus. Actuellement, la banque de données d'EST « dbEST » recense 8538 EST chez *E. gunnii*, 3951 EST chez *E. globulus*, 1574 EST chez *E. grandis* et 1131 EST chez *E. tereticornis* ce qui donne un total de 15194 séquences d'*Eucalyptus* disponibles pour la communauté scientifique. Plusieurs projets de recherche ont permis le séquençage et l'identification d'EST. Le projet Genolyptus au Brésil a isolé 135093 EST répartis dans 29 banques d'ADNc dérivés de divers tissus à différents stades de maturation : jeune feuilles, feuilles matures, fleur, racines, xylème, phloème. A l'inverse le projet FOREST (Brésil) axe une recherche plus ciblée sur des gènes impliqués dans la qualité du bois ou la résistance à des stress biotiques et abiotiques (Grattapaglia, 2004).

La combinaison du séquençage avec des banques d'EST spécifiques de tissus ou de conditions physiologiques particulières a ainsi menée à l'identification de nombreux gènes



d'intérêt (Hertzberg *et al.*, 2001). Des techniques d'analyse à haut-débit telles que les macro- ou microarrays (Kirst *et al.*, 2004) et l'hybridation soustractive suppressive (Diatchenko *et al.*, 1996) permettent l'analyse de profils d'expression de nombreux gènes dans un tissu donné. Ainsi, chez l'eucalyptus, Kirst *et al.* (2002) ont utilisé la technique de microarrays afin de comparer l'expression de gènes dans le xylème entre trois génotypes différents pour des caractères de qualité de bois et de croissance chez *E. globulus* et *E. grandis*. Encore chez l'eucalyptus, Paux *et al.* (2004) ont pu isoler des gènes préférentiellement exprimés dans le xylème par hybridation soustractive (SSH). Sur les 224 EST isolées, un tiers présentaient des homologies avec des protéines de fonction connue mettant en évidence la signalisation de l'auxine via la protéolyse ubiquitine-dépendante d'une part et des enzymes impliquées dans la biosynthèse et le remaniement des parois d'autre part.

Cependant, pour que ces gènes soient utilisés en sélection, il est nécessaire que leur variabilité moléculaire soit reliée à la variation du caractère ciblé. En améliorant des plantes, le sélectionneur ne s'intéresse qu'à certains caractères bien définis et cherche à sélectionner les meilleures combinaisons d'allèles des gènes d'intérêt correspondant. De plus, la majorité des caractères importants en sélection sont de type quantitatif, déterminés par de nombreux gènes. La stratégie la plus efficace pour identifier les gènes potentiellement intéressants pour la sélection passe par l'étude de la liaison entre des marqueurs moléculaires et la variation des caractères ciblés. Ces marqueurs doivent permettre la localisation de locus impliqués dans la variation de caractères quantitatifs (quantitative trait loci : QTL) d'intérêt comme la croissance et la qualité du bois. À terme, les marqueurs ont pour but de faciliter l'identification et la sélection des meilleurs génotypes (sélection assistée par marqueurs, Lefort *et al.*, 1990). L'identification de tels marqueurs suppose la création de cartes de liaison génétique. L'utilisation de ces cartes génétiques va permettre un balisage plus ou moins dense du génome afin de détecter des zones responsables de la variation de caractères qualitatifs et quantitatifs. Ces cartes ne visent pas à établir l'inventaire des gènes ou des séquences particulières, mais à marquer les portions du génome ayant un rôle majeur dans la variabilité de certains caractères. Chez les arbres, des cartes génétiques sont disponibles pour certaines essences, comme le peuplier (Bradshaw *et al.*, 1994), l'eucalyptus (Grattapaglia et Sederoff, 1994 ; Myburg *et al.*, 2003) et le pin (Devey *et al.*, 1994 ; Groover *et al.*, 1994).

Les premiers travaux de cartographie génétique et de recherche QTL, chez l'eucalyptus, ont été réalisés à l'aide de marqueurs dominants de type RAPD (Grattapaglia et Sederoff, 1994)

et AFLP (Marques *et al.*, 1998) et ont abouti à la construction de cartes génétiques saturés pour deux géniteurs d'un croisement interspécifique *E. urophylla* et *E. grandis*. Des QTL contrôlant la croissance des arbres, la forme du tronc et la densité du bois ont ensuite été localisés sur les cartes génétiques d'*Eucalyptus* (Verhaegen *et al.*, 1997), d'autres, reliés à la biosynthèse des lignines (Gion *et al.*, 2000) et à la propagation végétative (Grattapaglia *et al.*, 1995) ont été positionnés sur les cartes génétiques d'*E. grandis* et *E. urophylla*.

L'utilisation de marqueurs dominants ne permet pas d'étudier la sythénie entre espèces. De ce fait, de nombreux marqueurs codominants de type RFLP (Byrne *et al.*, 1995) et microsatellites (Brondani *et al.*, 1998) ont été ajoutés aux cartes d'origine et ont permis une avancée considérable dans la construction des cartes génétiques. Ces marqueurs codominants sont en effet très informatifs car multi-alléliques et présentent surtout l'avantage par rapport aux marqueurs dominants d'être positionnables sur les deux cartes parentales. La transférabilité des marqueurs microsatellites entre espèces appartenant au sous genre *Symphyomyrtus* a été étudiée par Byrne *et al.* (1996), Brondani et Grattapaglia (2002), Marques *et al.* (2002). Ces études ont révélées que 90% de ces marqueurs amplifiés chez *E. grandis* et *E. urophylla* (Brondani et Grattapaglia, 2002) et 78% de ces marqueurs développés chez *E. grandis* et *E. urophylla* pouvaient être transférés à *E. globulus* et *E. teriticornis* (Byrne *et al.*, 1996 ; Marques *et al.*, 2002). Cette transférabilité de marqueurs entre espèces suggère que l'établissement de cartes consensus est possible pour l'eucalyptus. A terme, une carte consensus va permettre de densifier des régions chromosomiques pauvres en marqueurs (ce qui sera très utile pour des espèces dont les parents présentent un bas niveau de polymorphisme), de valider l'ordre des marqueurs, d'étudier l'évolution du génome entre espèces, de prévoir la localisation d'un QTL par comparaison à une espèce pour laquelle un QTL a été identifié.

De ce fait, en 2006 les cartes génétiques d'*E. grandis* et *E. urophylla* ont été complétées avec l'ajout de 230 marqueurs microsatellites EMBRA par Brondani *et al.* (2006). Ces marqueurs ont permis une identification robuste des zones homologues entre les groupes de liaison des deux cartes, ainsi 82% des marqueurs avaient le même ordonnancement entre les deux cartes ce qui reflète une bonne conservation des génomes entre les espèces d'eucalyptus appartenant au sous genre *Symphyomyrtus*. Parmi ces marqueurs certains avaient déjà été positionnés sur deux cartes d'*E. globulus* établie par Bundock *et al.* (2000) et par Thamarus *et al.* (2002). Ainsi, la cartographie comparée entre les cartes d'*E. grandis*, *E. urophylla* et les



cartes d'*E. globulus* a permis de positionner des QTL identifiés chez *E. globulus* sur les cartes d'*E. grandis* et d'*E. urophylla*.

Identifier des QTL sur une carte génétique n'est pas suffisant pour à terme faire de la sélection assistée par marqueurs, il faut pouvoir cartographier un grand nombre d'EST sur les cartes génétiques afin d'observer une colocalisation de l'EST avec le QTL. Cette colocalisation permettra alors d'établir un lien entre la fonction putative de ces EST et le QTL ciblé. Les EST pour pouvoir être cartographiées doivent être variables dans le croisement étudié. Cette variabilité se caractérise par la présence de SNPs (mutations ponctuelles) ou d'INDEL (insertion ou délétion) au sein de la séquence. Différentes méthodes peuvent être utilisées pour révéler ce polymorphisme nucléotidique au sein d'un gène candidat. Certaines de ces méthodes sont basées d'une part sur les différences de stabilité de l'ADN : la SSCP (polymorphisme de conformation simple brin, Orita *et al.*, 1989), la DGGE (électrophorèse de l'ADN en gradient dénaturant ; Fisher et Lerman, 1979, 1983 ; Myers *et al.*, 1987), la TGGE (électrophorèse de l'ADN en gradient de température ; Rosenbaum et Riesner, 1987 ; Riesner *et al.*, 1991). D'autres méthodes plus récentes sont basées sur l'utilisation de la PCR en temps réel et de sondes fluorescentes spécifiques de zones polymorphes ciblées : Taqman (Mackay *et al.*, 2002) Molecular Beacon (Tyagi et Kramer, 1996) et Scorpion (Whitcombe *et al.*, 1999 ; Thelwell *et al.*, 2000 ; Mackay *et al.*, 2002).

Mon stage s'inscrit dans cette approche gène-candidat pour la qualité du bois, menée par le CIRAD. L'objectif est de cartographier des EST préférentiellement exprimées dans le xylème (isolées par SSH par Paux *et al.*, 2004) sur des cartes génétiques d'*E. urophylla* et *E. grandis* déjà établies par Gion (2001). Pour cela trois méthodes de génotypage ont été utilisées : la PCR en temps réel, la SSCP sur séquenceur LI-COR et la SSCP sur séquenceur ABI. Une analyse des séquences nucléotidiques des EST chez les 2 parents du croisement couplé à l'utilisation de la SSCP a permis de révéler les EST polymorphes. Ces EST ont ensuite été cartographiées par SSCP sur LI-COR et SSCP sur ABI sur les cartes génétiques d'*E. urophylla* et *E. grandis*.



Figure 1 : Aire naturelle des espèces *E. grandis* et *E. urophylla*. La localisation des provenances des deux géniteurs est indiquée en rouge.



## 2) **MATERIEL et METHODES**

### 2.1 - **Matériel végétal :**

Les hybrides utilisés pour la cartographie génétique proviennent d'un croisement interspécifique entre deux espèces du genre *Eucalyptus* : *E. urophylla* et *E. grandis*. Ces deux espèces appartiennent à la section *Transversaria* du sous genre *Symphyomyrtus*. Le géniteur *E. urophylla* (individu 14.144), utilisé comme parent femelle provient du Monte Lewotobi de l'île de Florès située au nord-ouest de l'Australie. Le géniteur *E. grandis* (individu 9.21), utilisé comme parent mâle provient d'Atherton, région de la Cordillère australienne située au Queensland en Australie (Figure 1).

Le croisement *E. urophylla* x *E. grandis* a été effectué en 1992 dans la région côtière de Pointe-Noire (Congo) et totalise 201 individus pleins-frères. Cette descendance correspond au croisement de référence du CIRAD pour les travaux de cartographie génétique et de détection de QTL (Verhaegen et Plomion, 1996 ; Verhaegen *et al.*, 1997 ; Gion *et al.*, 2000).

#### **Extraction d'ADN génomique :**

L'extraction d'ADN génomique a été réalisée à partir de feuilles séchées suivant le protocole de Saghai-Marouf *et al.* (1984). Pour chaque individu, 100 mg de feuilles séchées sont broyées en présence d'azote liquide. La poudre obtenue est incubée 1h à 65°C en présence de 1,3 ml de tampon d'extraction stérile (NaCl 1,4M, Tris HCL 0,1M pH8, EDTA 0,02M, ATMA B 20g/L), 0,1 % de mercaptoéthanol et 1% de polyvinylpyrrolidone. Après refroidissement des tubes à l'air libre, 400 µl de dichlorométhane sont ajoutés. Les extraits sont mis en agitation douce pendant 10 min à 3000 g (6000T/min). Le surnageant est prélevé et cette étape est effectuée 2 fois. L'ADN génomique est alors précipité à froid par ajout de 500 µl d'isopropanol glacé et de 7 µl de NaCl 5M. Les extraits sont placés 1h à -20°C (ou une nuit à -4°C) puis centrifugés 10 min à 3000 g. Le surnageant est éliminé et les culots sont rincés avec 500 µl d'éthanol à 76°, séchés sous vide au speed-vac 30 min à 40°C et remis en suspension dans 100µl d'eau stérile à 65°C (bain-marie). Si le culot d'ADN est important, il est repris dans 200 à 500 µl d'eau. Les produits d'extraction sont stockés à -20°C.

## 2.1 - Matériel végétal :

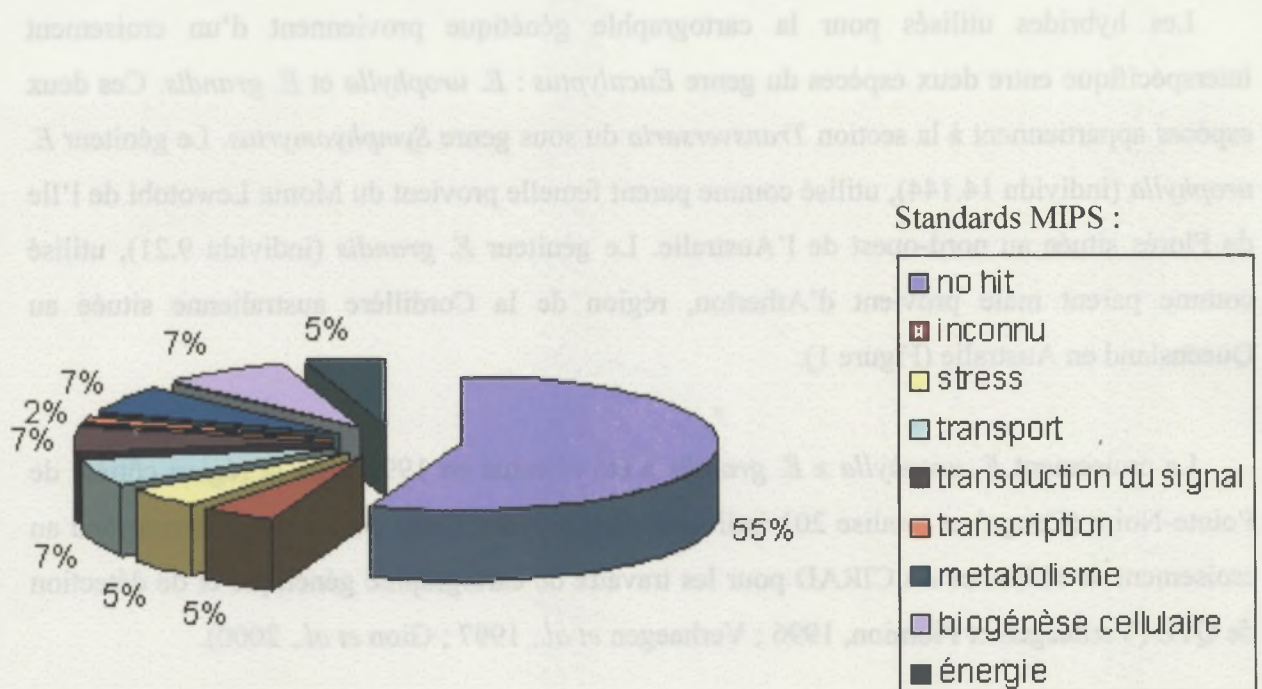


Figure 2 : Catégories fonctionnelles des 61 EST obtenues chez *E. gunnii* en fonction des standards MIPS.



Les produits d'extraction ont été dosés au spectrophotomètre NANODROP ND-1000. Les solutions mères ont alors été diluées à 5 ng/μl avec de l'eau stérile. Les solutions de travail obtenues sont stockées à -20°C.

## **2.2 - Caractéristiques des EST étudiées :**

Les séquences de gènes utilisées sont des EST (Expressed sequence tag) développées chez *E. gunnii* par le CNRS de Toulouse (Paux, 2004). Ces gènes surexprimés dans le xylème chez *E. gunnii* ont été isolés par les techniques d'hybridation soustractive suppressive (SHH). Les techniques de cDNA array et RT-PCR ont permis l'analyse de leurs niveaux d'expression. De cette étude 224 EST indépendantes ont été identifiées. Ces EST correspondent à différentes catégories fonctionnelles en fonction des standards MIPS (inconnu, métabolisme, énergie, croissance cellulaire et division, transcription, métabolisme, transport, biogénèse cellulaire, transduction du signal, stress).

61 EST provenant d'un sous-échantillonnage de ces 224 EST sont utilisées pour mon étude. Pour les 61 EST étudiées, la fonction, leurs niveaux d'expression dans le xylème et leur taille sont représentés dans l'annexe 1. Ces EST correspondent en majorité à des extrémités 3' de gènes et leur taille est comprise entre 137 et 769 pb.

La répartition des EST par rapport à leur fonction probable est représentée en Figure 2. Parmi ces EST 55% ne présentent aucune homologie avec des gènes connus, 5% présentent des homologies avec des gènes de fonction inconnue et 40% présentent des homologies avec des gènes de fonction connue. Ces dernières comme illustré dans l'annexe 1 sont pour la plupart directement impliquées dans la synthèse des parois (ex : 12C02 codant une protéine arabinogalactane), d'autres ont des fonctions plus classiques et interviennent dans la transduction d'un signal (12F09), dans des phénomènes de stress (09C04) ou dans la voie de la protéolyse ubiquitine-dépendante permettant de réguler la transcription de l'auxine (EST 07E09) (Joazeiro and Weissman, 2000).

### **2.3 - Amplification des EST par PCR dirigée :**

Les paires d'amorces spécifiques de chaque EST ont été définies avec le logiciel OLIGO 4.0 (Rychlik *et al.*, 1990). Pour chacune des EST, les séquences des amorces sont spécifiques à la séquence disponible d'*E. gunnii* et ont été préalablement testées sur *E. globulus*. Ces amorces ont une taille d'environ 20 nucléotides. Afin de faciliter la mise en évidence du polymorphisme, les amorces ont été définies de manière à amplifier des fragments d'au minimum 100 pb.

Comme indiqué dans l'annexe 2, les températures d'hybridation des amorces sont comprises entre 48 et 68°C et les fragments amplifiés ont une taille attendue située entre 100 paires de bases (pb) (EST 06C05) et 521 pb (EST 06E02).

Les réactions PCR sont réalisées sur un thermocycleur de type TETRAD2 MJR avec des plaques ABgene de 96 puits. Par puits : le volume final est de 20 µl avec 4 µL d'ADN génomique à 5 ng/µl, 2 µl de chaque amorce à 2 µM, 0,16 µl de Taq DNA polymérase à 5U/µl et 10 µl de tampon 2X (contenant 80 µl de MgCl<sup>2</sup> à 50 mM, 80 µl de dNTPs à 5 mM, 200 µl de Tp10X, 640 µl H<sub>2</sub>O).

Les conditions d'amplification sont les suivantes : une dénaturation initiale de 4 min à 94°C puis 35 cycles avec une étape de dénaturation (30 sec à 94°C), une étape d'hybridation (1 min à une température spécifique de chaque EST) et une étape d'extension d'1 min à 72°C. Cette première partie s'achève par une élongation de 5 min à 72°C.

L'amplification spécifique des EST est vérifiée par autoradiographie après une électrophorèse sur un gel d'agarose 2% immergé dans du TBE 0.5X et coloré au bromure d'éthidium (BET).

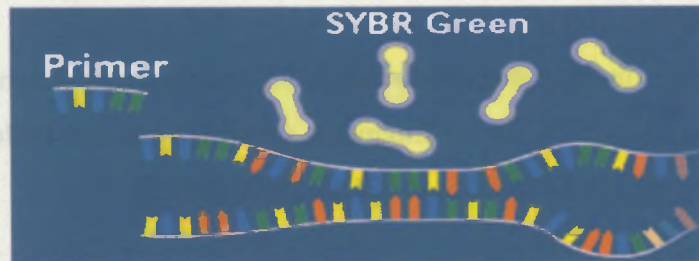
### **2.4 - Mise en évidence du polymorphisme :**

La stratégie suivie pour la mise en évidence du polymorphisme au sein de la descendance pour les 61 EST étudiées a été :

-1) le séquençage des EST sans étape préalable de clonage chez les deux géniteurs afin de confirmer par BLASTn la nature des fragments amplifiés et de révéler ou non une



1) SYBR Green libre, non fluorescent :



2) Début de l'élongation, incorporation du SYBR Green dans l'ADN db :



3) Fluorescence du SYBR Green lors de son incorporation à l'ADN db :

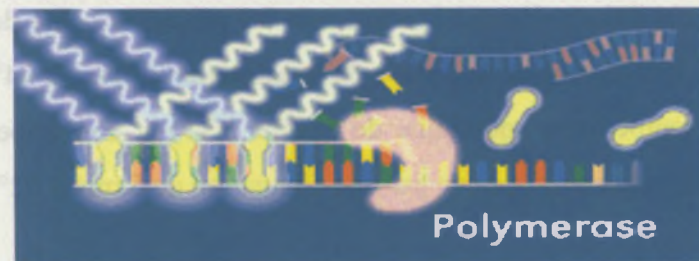


Figure 3 : Illustration de Martin-Laurent *et al.*, (2004) représentant la fluorescence du SYBR Green lors de son incorporation dans l'ADN double brin.

hétérozygotie (présence de deux allèles différents) chez les deux géniteurs *E. grandis* et *E. urophylla*.

-2) la sélection des EST polymorphes chez au moins un des deux parents par SSCP sur LICOR.

-3) l'amplification des EST polymorphes sélectionnées et le génotypage des descendants par trois techniques différentes : la PCR en temps réel, la SSCP sur séquenceur LI-COR et la SSCP sur séquenceur ABI.

#### **2.4.1 - Séquençage :**

Le séquençage concerne la détermination de la séquence des gènes c'est-à-dire l'enchaînement précis des nucléotides du fragment ADN étudié. La méthode la plus utilisée actuellement est la méthode de Sanger. Cette technique utilise la réaction de polymérisation de l'ADN avec la «Taq Polymérase», les quatre désoxynucléotides (dATP, dCTP, dGTP, dTTP) et en plus faible quantité, quatre classes de didésoxyribonucléotides (ddATP, ddCTP, ddGTP, ddTTP) marqués par des fluorochromes différents et se différenciant des dNTP par leur absence de groupement hydroxyle OH sur l'extrémité 3'. Ces ddNTP une fois incorporés à la nouvelle chaîne synthétisée, stoppent la poursuite de l'élongation car l'absence de groupement OH à leur extrémité 3' empêche la fixation des dNTP par la Taq polymérase en aval du néo-brin. Il en résulte la formation de fragments d'ADN de tailles variables.

Ces fragments d'ADN sont analysés sur séquenceur à capillaire MegaBace. Le protocole utilisé est détaillé en annexe 3.

#### **2.4.2 - Révélation du polymorphisme par PCR en temps réel :**

Cette technique consiste en l'analyse de la température de fusion d'un individu grâce à l'incorporation du fluorochrome SYBR green (Le Morvan *et al.*, 2005). La température de fusion ou T<sub>m</sub> (melting temperature) est la température à laquelle 50% des ADN sont sous forme désappariée (sous forme simple brin) lors d'une dénaturation des produit PCR. Le fluorochrome SYBR green est capable de se lier à l'ADN double brin (ADNdb). Comme illustré en Figure 3, le SYBR Green en solution n'est pas fluorescent sous sa forme libre, alors qu'il est très fluorescent lorsqu'il est lié à l'ADNdb. Une PCR pour l'EST d'intérêt est donc réalisée en présence du SYBR green (qui va s'incorporer à l'ADNdb durant l'étape



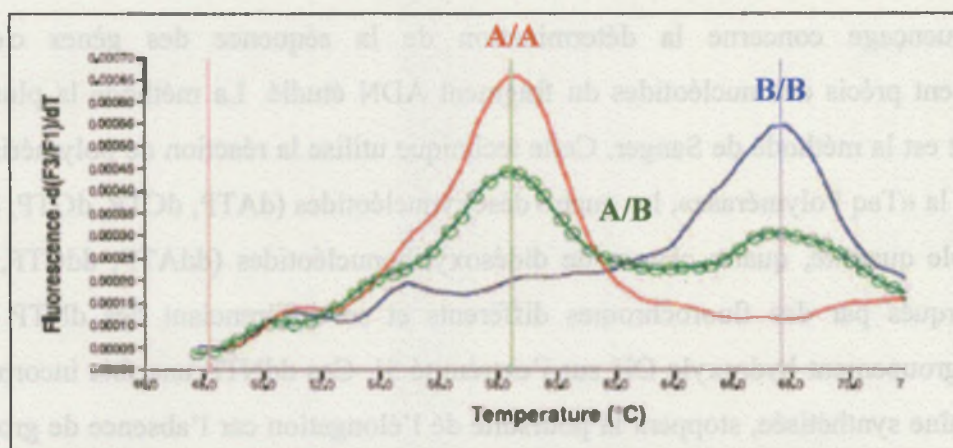


Figure 4 : Courbes de fusion d'individus : homozygote pour l'allèle A (rouge), homozygote pour l'allèle B (bleu) et hétérozygote AB (vert). (d'après Steffensen *et al.*, 2002).

d'élongation), suivie d'une étape de Melting curve (montée de température de 65°C à 95°C). Cette montée en température permettra la séparation de l'ADNdb et aura pour conséquence une chute de la fluorescence émise suite à la libération du fluorochrome. Quand 50% des ADN double brin sont dissociés, la fluorescence chute brutalement, c'est à cette température que correspond la température de fusion du produit synthétisé. On obtient une valeur précise du  $T_m$  en effectuant une courbe de dérivée négative de la chute de fluorescence en fonction du temps :  $-d(UFR)/dt$ . UFR signifiant : unité relative de fluorescence. Le  $T_m$  sera ainsi représenté par un pic (Figure 4).

Sachant que chaque amplicon devrait théoriquement avoir son propre  $T_m$ , il doit être possible d'analyser la ségrégation allélique d'une descendance en comparant le  $T_m$  de chaque descendant avec ceux de leurs géniteurs 9.21 et 14.144. A titre d'exemple la Figure 4 illustre les  $T_m$  de deux parents homozygotes (A/A et B/B) et d'un descendant hétérozygote A/B dont le  $T_m$  représenté par un double pic recoupe les  $T_m$  parentaux. L'*Eucalyptus* étant une espèce diploïde on s'attend à observer deux  $T_m$  par individu hétérozygote c'est-à-dire un  $T_m$  spécifique à chaque allèle.

#### ▪ **Protocole :**

L'automate de PCR (PTC-200) possède un bloc classique assurant les cycles PCR et un bloc de détection capable de mesurer l'intensité de fluorescence émise. Lorsque l'amplification atteint un seuil supérieur au bruit de fond, la fluorescence émise est directement corrélée à la quantité d'ADN.

La PCR en temps réel nécessite l'utilisation d'un kit spécifique appelé Bio-Rad IQ SYBR Green Supermix contenant un tampon de réaction (KCL 100 mM, Tris HCL 40 mM, pH=8.4), des dNTP (400  $\mu$ M de chaque), une Taq DNA polymérase, du  $MgCl_2$  (6mM). La réaction PCR est réalisée dans un volume final de 25  $\mu$ l contenant 0,01  $\mu$ M de chaque amorce, 50 ng d'ADN et 12,5 de MIX Bio-Rad (Karsai *et al.*, 2002).

Le programme utilisé se divise en deux phases. La première phase est une amplification classique. Elle comprend une dénaturation initiale de 3 min à 95°C puis 35 cycles avec une étape de dénaturation (30 sec à 94°C), une étape d'hybridation (1 min à une température spécifique de chaque EST et une étape d'extension (1 min à 72 °C). Cette première partie s'achève par une élongation de 5 min à 72°C. La seconde partie permet la détermination de la température de fusion. Il s'agit de réaliser des tests de dénaturation des doubles brins néo-



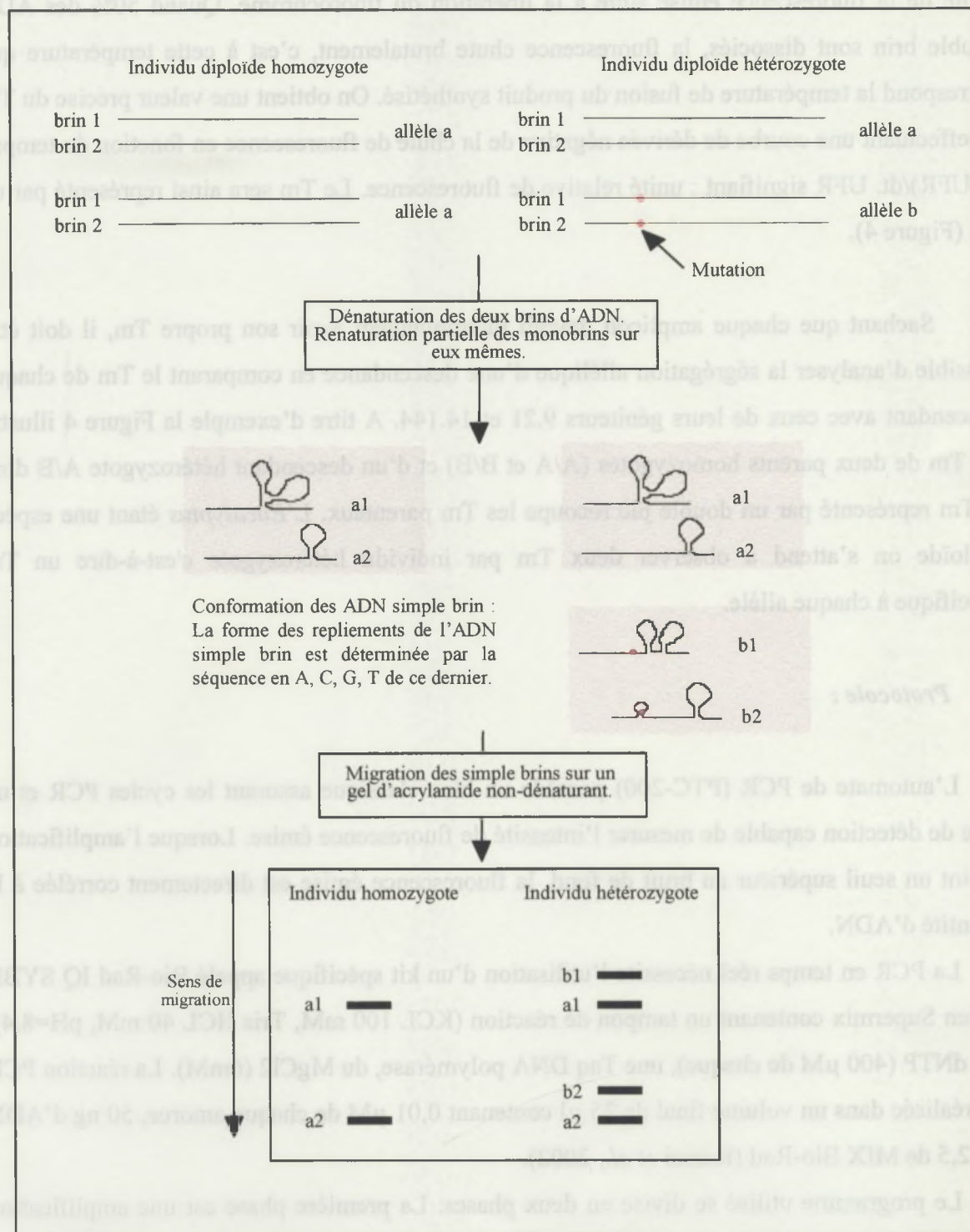


Figure 5 : Révélation du polymorphisme par la technique SSCP (Single Strand Conformation Polymorphism). D'après Gion (2001).

synthétisés en passant de 65°C à 95°C avec des incréments de 0,1°C toutes les 10 secondes (Karsai *et al.*, 2002). Le T<sub>m</sub> est ensuite déterminé à partir de l'évolution de la fluorescence émise dans le temps en traçant la courbe  $-d(\text{URF})/dT$ . Le T<sub>m</sub> est la valeur de la température pour laquelle cette courbe a atteint son maximum.

### **2.4.3 - Révélation du polymorphisme par la technique SSCP :**

La technique SSCP (Single Strand Conformation Polymorphism) développée par Orita *et al.* (1989) permet de détecter un polymorphisme de conformation d'ADN simple brin par électrophorèse sur gel non dénaturant (Figure 5). Chaque ADN simple brin, après dénaturation, va se réassocier partiellement pour former des structures secondaires et cela en fonction de la séquence des fragments considérés. L'électrophorèse va permettre de séparer différents fragments selon les conformations obtenues. La substitution d'une base au niveau de la molécule d'ADN donne des conformations différentes qui migreront différemment sur un gel de polyacrylamide en conditions non dénaturantes. Les marqueurs ainsi générés sont de type codominants et multialléliques.

Très sensible sur des petits fragments compris entre 150 et 300 pb, la technique ne permet pas de révéler de façon exhaustive tout le polymorphisme existant pour des fragments plus longs. Pour augmenter la précision et l'information, il est alors possible de l'appliquer à des fragments préalablement digérés ou de faire varier des conditions de migration (température, composition de gel avec ajout de glycérol).

En SSCP classique, les ADN sont marqués au bleu de bromophénol et xylène de cyanol et le repérage de l'ADN sur le gel d'acrylamide s'effectue par une coloration des gels au nitrate d'argent réalisée selon la méthode de Creste *et al.* (2001).

#### **2.4.3.1 - La SSCP sur LI-COR :**

Dans le cas d'une SSCP sur LI-COR, la détection des fragments ADN sur gel d'électrophorèse, se fait par un marquage fluorescent de ces fragments (Schuelke, 2000). De ce fait l'amorce forward portera une extension de séquence appelée M13 qui permettra la fixation du fluorochrome choisi : IRD700 (longueur d'excitation de 700 nm) ou 800 (longueur d'excitation de 800nm). Les produits PCR sont dénaturés avant d'être analysés sur séquenceur LI-COR. Dans le séquenceur les fragments d'ADN subissent une électrophorèse sur gel



d'acrylamide. La révélation de la fluorescence se fait par excitation laser du fluorochrome à une longueur d'onde d'excitation qui lui est spécifique.

▪ **Protocole :**

Une première PCR est réalisée avec 2 µL d'ADN génomique à 5 ng/µl, 1 µl de chaque amorce (amorce sens avec extension M13 et amorce antisens sans extension), 0,08 µl de Taq DNA polymérase à 5U/µl et 5 µl de tampon 2X (contenant 80 µl de MgCl<sub>2</sub> à 50 mM, 80 µl de dNTPs à 5 mM, 200 µl de Tp10X, 640 µl H<sub>2</sub>O).

Les conditions d'amplification sont les suivantes : une dénaturation de 4 minutes à 94°C, suivie de 35 cycles comprenant une dénaturation de 30 secondes à 94°C, une hybridation d'une minute à une température spécifique de chaque EST, une élongation d'une minute à 72°C. A tout cela s'ajoute une élongation finale de 5 minutes à 72°C.

Pour chaque produit PCR, la spécificité de l'amplification est vérifiée par une électrophorèse sur gel d'agarose à 2% dans du TBE 0,5X et une coloration au bromure d'éthidium (BET) visualisable par autoradiographie.

Une seconde PCR est réalisée avec 1,5 µl de produit PCR, 0,04 µM de chaque amorce (amorce M13 marquée à l'IRD700 ou IRD800 et amorce antisens), 5 U Taq, 1,5 µl Tampon 10X (Invitrogene), 2 mM MgCl<sub>2</sub>, 0,2 mM dNTPs et qsp 20 µl d'eau MilliQ (Schuelke, 2000). Les conditions d'amplification sont les suivantes : une dénaturation de 4 minutes à 94°C, suivie de 35 cycles comprenant une dénaturation de 30 secondes à 94°C, une hybridation d'une minute à 55°C (T<sub>m</sub> spécifique de l'IRD700 et 800), une élongation d'une minute à 72°C et enfin une élongation finale de 5 minutes à 72°C.

Pour la révélation sur LI-COR, 1 µl de produit PCR sont ajoutés à 3 µl de bleu (78 ml formamide, 10 ml de xylène cyanol 1%, 10 ml de bleu de bromophénol 1%, 2 ml d'EDTA [0,5 M] pH 8). Les échantillons sont dénaturés 3 minutes à 94°C puis aussitôt refroidis dans la glace. Ils sont ensuite déposés sur un gel d'acrylamide non dénaturant (5,23 ml MDE, 1,27 ml TBE 10X, 14,40 ml H<sub>2</sub>O, 13,75 ml TEMED, 137,50 µl d'APS 10%). Les conditions de migration électrophorétique testées ont été : un voltage de 1000, un courant de 40 mA, une puissance de 40W et une température de 25°C pendant un temps de migration de 12h30. La révélation de la fluorescence se fait par excitation laser à une longueur d'onde de 700 nm pour l'IRD700 et 800 nm pour l'IRD800.

#### **2.4.3.2 - SSCP sur séquenceur ABI :**

Les produits PCR sont dénaturés avant d'être analysés sur séquenceur ABI. Dans le séquenceur les fragments vont subir une migration. Durant leur électrophorèse les fragments d'ADN amplifiés passent devant une fenêtre de détection qui identifie le type de fluorochrome associé à l'amorce «forward» (amorce sens). Un laser va exciter le fluorochrome qui va émettre à une certaine longueur d'onde détectable par la caméra CCD de la fenêtre de détection. Le marqueur de taille ajouté aux produits PCR avant la migration permet de relier le temps de migration des fragments amplifiés à leur taille exprimée en nombre de paires de base et facilite ainsi l'interprétation des résultats. Les amorces plus petites que les produits PCR passent en premier devant la cellule car leur migration, facilitée par leur petite taille, est plus rapide (Tozaki, 2002).

Le signal perçu par la caméra CCD apparaît à l'écran sous forme de pics correspondants aux différentes structures secondaires adoptées par les brins d'ADN.

##### **▪ Protocole :**

Le principe est le même que pour une SSCP sur LI-COR. L'amorce forward porte ainsi une extension M13 qui permettra la fixation du fluorochrome choisi : FAM6 au lieu de l'IRD700 ou 800 (Schuelke, 2000). Les PCR 1 et 2 sont identiques à celles utilisées en SSCP sur LI-COR (protocole 2.4.3.1). Après PCR : 1 µL des produits PCR est mélangé à 0,5 µl de marqueur de taille LIZ600 (adapté à des fragments de taille comprise entre 10 et 600pb) et 10,5 µl de HIDi formamide (dénaturant qui va aider à la conservation des structures secondaires formées après dénaturation) dans une plaque AB gene 1100. Les échantillons sont alors dénaturés 3 minutes à 94°C puis aussitôt refroidis dans la glace.

Après dénaturation, les échantillons sont analysés sur un séquenceur ABI 3130 XL (Applied Biosystems, Foster City, CA, USA) de 16 capillaires (36 cm). Dans un premier temps les échantillons sont injectés électro-cinétiquement (par différence de potentiel) dans les capillaires avec un temps d'injection de 20 secondes à 15 kV. L'électrophorèse dans le gel CAP polymère 5 % durera 45 minutes à une température de 25°C.

Les résultats sont alors observés grâce au logiciel GeneMapper.



## 2.5 - Cartographie génétique des EST polymorphes :

Les cartes génétiques utilisées lors de cette étude ont été construites avec des marqueurs dominants de type RAPD (Verhaegen et Plomion, 1996) selon une stratégie de double pseudo-testcross (appliquée pour la première fois chez l'*Eucalyptus* par Grattapaglia et Sederoff, 1994). Cette stratégie, basée sur l'utilisation de marqueurs dominants hétérozygotes pour l'un des parents et ségrégeant dans la descendance dans les proportions mendéliennes  $\frac{1}{2} : \frac{1}{2}$ , permet la construction de cartes génétiques spécifiques à chacun des parents. On peut ensuite identifier les groupes de liaisons homologues entre les deux cartes à l'aide de marqueurs communs hétérozygotes chez les deux parents qui servent de marqueurs «ponts».

Les marqueurs codominants SSCP présentent jusqu'à quatre allèles différents qui ont été lus 1, 2, 3 et 4. Il est nécessaire de recoder cette information en 0/1 afin d'établir leur lien avec les marqueurs dominants déjà présents sur les cartes. Ainsi, nous assignons un code pour chaque allèle de chaque parent et recodons la descendance en conséquence, on lit donc ces marqueurs codominants comme des marqueurs dominants.

Avant d'entreprendre l'estimation de la liaison entre marqueurs, un test statistique de  $\chi^2$  est appliqué au niveau de chaque marqueur afin de comparer les effectifs observés des classes de génotypes aux effectifs théoriques, sous l'hypothèse d'une ségrégation mendélienne. Ainsi dans le cas où un seul des parents est hétérozygote, la ségrégation observée doit être de type  $\frac{1}{2} : \frac{1}{2}$ . Le positionnement de l'EST sur la carte génétique ne sera alors possible que chez ce parent hétérozygote. Dans le cas où les deux parents sont hétérozygotes, la ségrégation attendue doit être de type  $\frac{1}{4} : \frac{1}{4} : \frac{1}{4} : \frac{1}{4}$ . L'EST pourra alors être cartographiée chez les deux parents et permettra d'identifier des zones homologues entre les cartes génétiques des deux parents.

La liaison génétique entre deux marqueurs est estimée à partir de la fréquence des recombinaisons méiotiques. Le taux de recombinaison a pour valeur minimale 0, et pour valeur maximale 0,5 qui signifie qu'il n'y a pas de liaison (indépendance) entre les deux marqueurs. Un test de liaison est effectué à l'aide du rapport de vraisemblance ou LOD Score défini par :

$$\text{LOD}(\theta) = \log_{10} [L(\theta) / L(\frac{1}{2})]$$

où  $L(\theta)$  représente la vraisemblance d'un taux de recombinaison compris entre 0 et 0,5 et  $L(1/2)$ , la vraisemblance d'une absence de liaison entre les deux marqueurs. Un LOD de 3 signifie que la liaison entre les marqueurs est 1000 fois plus probable que l'indépendance.

Les EST sont positionnées sur des cartes génétiques d'*E. grandis* et *E. urophylla* à l'aide du logiciel MAPMAKER. La première étape consiste à établir des groupes de liaison par la fonction «GROUP» en fixant un LOD minimum et une fréquence de recombinaison maximale. Dans notre étude une valeur de LOD de 5 et une fréquence de recombinaison de 0,30 ont été fixées. La seconde étape consiste à identifier notre EST dans un groupe de liaison. Une fois l'identification réalisée, la fonction «FIRST ORDER» va permettre d'agencer les marqueurs entre eux dans l'ordre le plus probable. La fonction «RIPPLE» permettra alors d'affiner localement cet ordre, en réalisant des permutations de triplets de marqueurs adjacents et calculera combien de fois un ordre donné est plus «probable» que celui obtenu par la fonction «FIRST ORDER».

Afin de déterminer les distances génétiques séparant les différents marqueurs, la fonction Kosambi est choisie car elle tient compte de l'interférence génétique (contraintes mécaniques lors de la méiose qui empêchent la formation de double crossing-over entre 2 marqueurs).



### 3) RESULTATS

#### 3.1 - Amplification des EST chez les géniteurs 9.21 et 14.144 :

Les tests d'amplification pour les deux géniteurs du croisement *E. urophylla* x *E. grandis* ont été réalisés avec une température d'hybridation des amorces optimisée chez *E. globulus*. Sur les 61 EST testées, 41 couples d'amorces ont donné un produit d'amplification chez les 2 géniteurs 9.21 (*E. grandis*) et 14.144 (*E. urophylla*) (annexe 2).

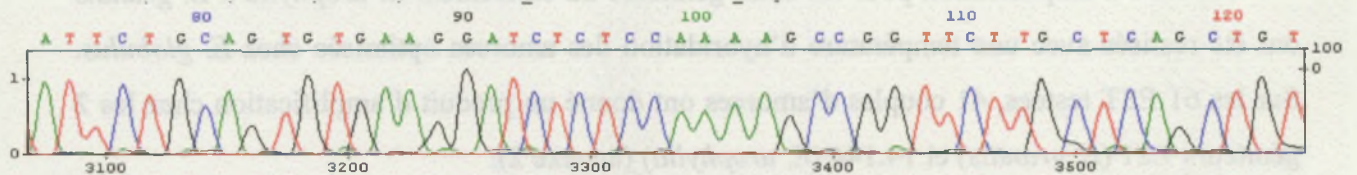
Parmi les 42 produits d'amplification : 27 ont une taille observée identique à la taille attendue chez *E. gunnii*, 14 ont une taille supérieure à la taille attendue laissant supposer qu'un intron ait été amplifié. Enfin, pour l'EST 12B02 la taille observée est inférieure à la taille attendue ce qui laisse penser qu'il y a eu une amplification non spécifique.

Les 19 autres couples d'amorces n'ont pas permis d'obtenir un produit d'amplification pour les deux espèces ou seulement chez un parent pour les conditions d'amplification testées. Sur ces 19 couples, quatre produits d'amplification (EST 07B07, 08C03, 08C06 et 12G04) présentent des tailles observées supérieures à la taille attendue. Une optimisation des conditions d'amplification (gradient pour la température d'hybridation) devrait permettre l'amplification de ces EST chez les deux géniteurs. La taille moyenne des introns est de 165 pb avec une taille minimale de 11 pb pour l'EST 08C03 et une taille maximale de 674 pb pour l'EST 12G04.

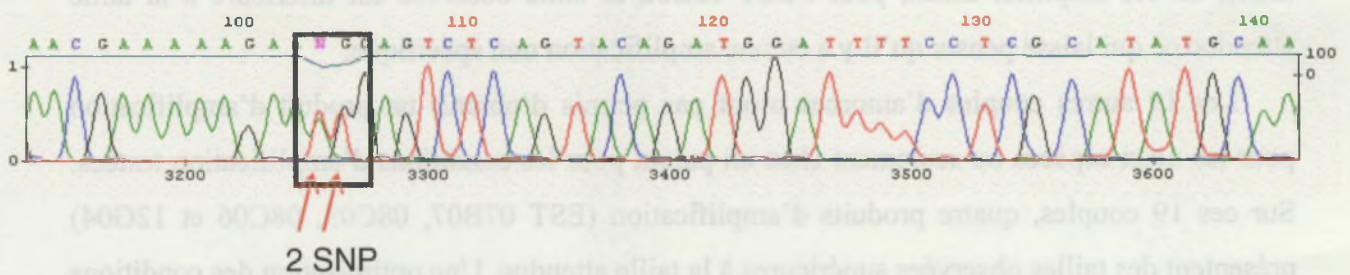
#### 3.2 – Mise en évidence de la variabilité chez les parents :

Les EST possédant un produit d'amplification unique (une seule bande sur le gel d'agarose 2%) ont été séquencées chez les 2 géniteurs *E. grandis* et *E. urophylla*. Afin de vérifier la nature du fragment amplifié la séquence nucléotidique a été comparée aux séquences disponibles dans les banques de données (programme BLASTn), avec des paramètres par défaut proposés par NCBI. Pour *E. grandis* et *E. urophylla*, les analyses BLASTn des EST amplifiées ont révélées que celle-ci correspondent bien aux EST d'*E. gunnii* confirmant ainsi la nature des fragments amplifiés. Ainsi l'EST 12BO2 pour laquelle la taille observée est inférieure à la taille attendue correspond bien à une amplification spécifique.

A) Electrophorégramme de 07B07 chez *E. grandis* :



B) Electrophorégramme de 08F11 chez *E. grandis* :



C) Electrophorégramme de 10E08 chez *E. grandis* :

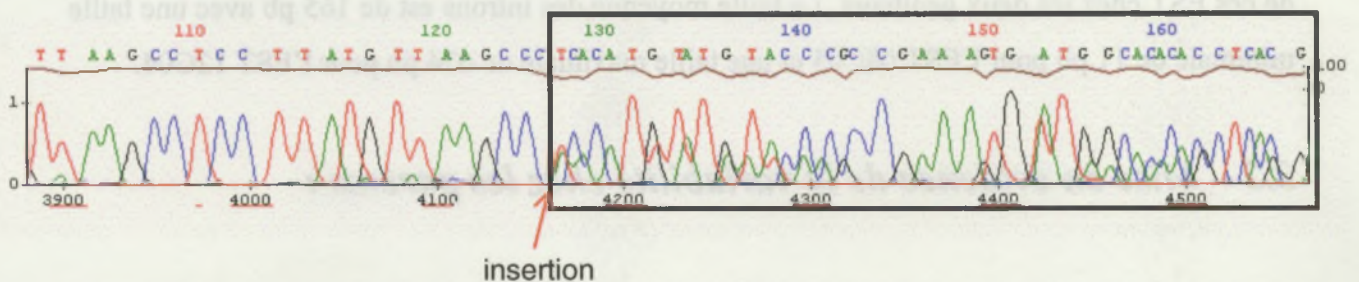


Figure 6 : Electrophorégrammes de séquences sans variation (A), présentant des SNP (B), ou un INDEL (C).



Le séquençage des produits d'amplification a aussi permis de mettre en évidence la variabilité au sein des fragments EST chez *E. grandis* et *E. urophylla*. En effet, la lecture des profils de séquençage permet de détecter les variations nucléotidiques ponctuelles (SNP, single nucleotide polymorphism) et les insertion/délétion (INDEL). Certaines EST ne présentent aucune variation à la lecture de l'électrophorégramme, comme pour l'EST 07B07 séquencée chez *E. grandis* illustrée par la Figure 6-A. Cette absence de variation peut s'expliquer par :

- l'homozygotie des génotypes parentaux (présence de deux allèles identiques pour chaque parent). Dans ce cas, l'EST sera monomorphe dans la descendance.
- l'absence d'amplification d'un des allèles. On obtient alors un produit d'amplification correspondant à un seul allèle. Dans ce cas, l'analyse de ségrégation au sein de la descendance révélera malgré tout un polymorphisme pour l'EST.

La simple lecture des électrophorégrammes ne nous permet pas de conclure définitivement sur l'hétérozygotie ou l'homozygotie des génotypes parentaux. Le séquençage offre tout de même une information assez précise sur le niveau et le type de variabilité au sein des séquences.

En général, pour les EST variables chez les parents du croisement, deux types de variabilité ont pu être mise en évidence grâce à la lecture des électrophorégrammes :

- une variabilité de type SNP (mutation ponctuelle au niveau des nucléotides) qui va se traduire au niveau de la séquence par la superposition de deux pics. Ainsi, comme illustrés sur la Figure 6-B (cas de l'EST 08F11 séquencée chez *E. grandis*), les deux allèles parentaux se distinguent par la présence de deux SNP (A/T et T/G).
- une variabilité de type INDEL. Une insertion d'un nucléotide sur un des allèles représentée en Figure 6-C (cas de l'EST 10E08 séquencée chez *E. grandis*) provoque un décalage des nucléotides ce qui se traduit par une superposition des pics au niveau des électrophorégrammes. La délétion d'un nucléotide sur un des allèles donne le même type de profils.

L'annexe 4 présente les types de variabilités observées (SNP, INDEL) pour chaque EST chez *E. grandis* et *E. urophylla*. Chez les deux géniteurs seulement la moitié des électrophorégrammes ont été exploitables. La mauvaise qualité de résultats est due à une purification trop longue qui a provoqué une perte partielle des produits PCR. Cependant, les séquences exploitables ont permis pour la suite des travaux de valider les résultats de polymorphisme obtenus par Q PCR et SSCP sur LI-COR.



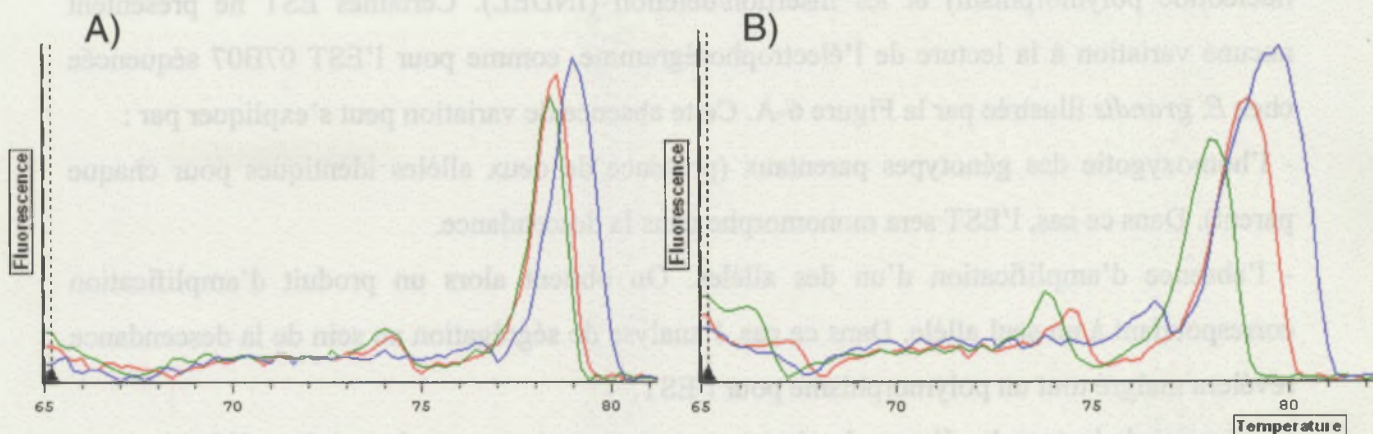


Figure 7 : Evolution de la fluorescence en fonction de la température pour l'EST 09H06 chez le parent *E. urophylla* (En rouge), le parent *E. grandis* (en bleu) et un hybride (en vert). Les courbes de fusions sont données sans DMSO (A) et avec DMSO (B).

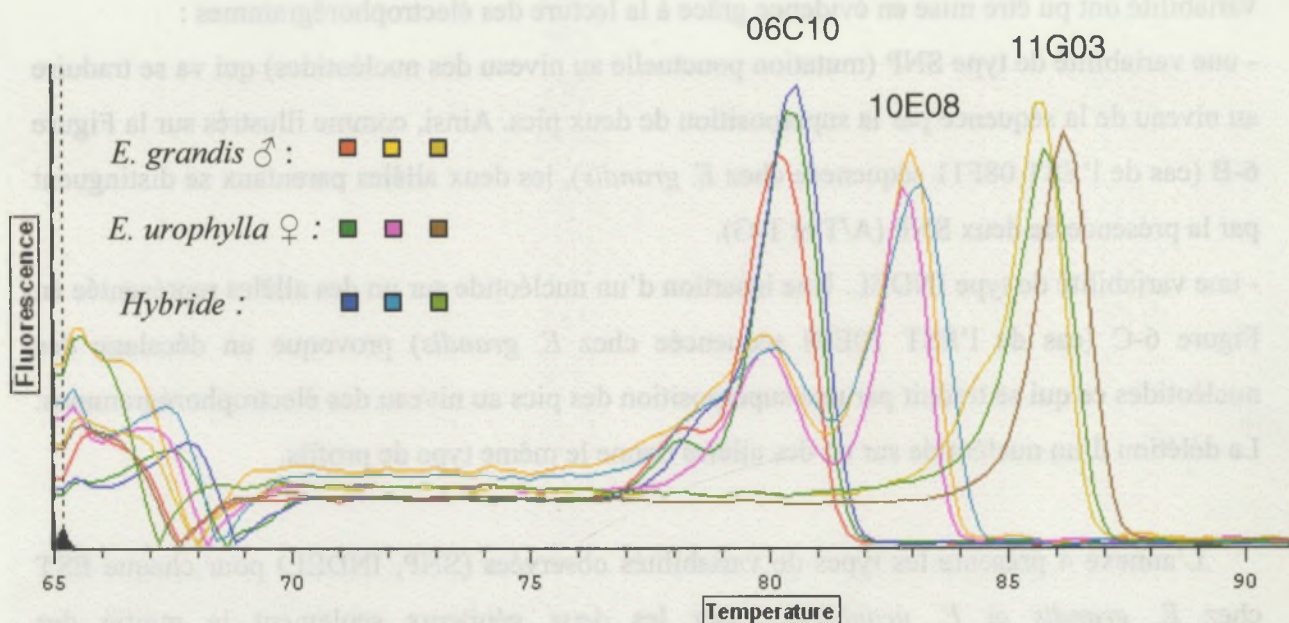


Figure 8 : Courbes de fusion pour trois EST (06C10, 10E08 et 11G03) pour les deux parents du croisement et un hybride. Ces trois EST sont polymorphes chez au moins un des parents.



### 3.3 - Mise en évidence du polymorphisme chez les descendants :

Trois méthodes ont été utilisées pour mettre en évidence le polymorphisme au sein de la descendance hybride : la PCR en temps réel ou PCR quantitative, la SSCP sur séquenceur LICOR et la SSCP sur séquenceur ABI.

#### 3.3.1 - La technique PCR en temps réel pour le génotypage haut débit :

La PCR en temps réel est une technique peu utilisée dans le génotypage et préconise en règle générale l'utilisation d'amorces marquées spécifiques des régions polymorphes (sondes Taqman, Molecular beacon, Scorpion) (Poitras *et al.*, 2002) Notre étude est simplement basée sur l'analyse des Tm des EST et aucune amorce spécifique des zones mutées n'a été définie.

Pour cela une phase d'optimisation était nécessaire. Des EST (06C10, 10E08, 11E10, 11F03 et 11G03) provenant de mon étude et l'EST (09H06) cartographiée par Carrouché (2003) ont été testées pour les parents 9.21 et 14.144 et l'hybride 101. Elles ont été choisies de manière à représenter les différents types de variabilité observée (SNP, INDEL ou monomorphe) au sein d'une séquence (annexe 5).

Dans notre étude l'utilisation de SYBR green pur à la place du KIT biorad a donné des profils peu lisibles avec un bruit de fond important. Le KIT BIORAD a permis d'obtenir des profils plus lisibles (Figure 7-A). On observe ainsi pour l'EST 09H06 un seul Tm par individu aux alentours de 78°C, il est donc difficile d'effectuer une lecture du polymorphisme. Comme illustré sur la Figure 7-B, l'ajout du dénaturant DMSO (Frackman *et al.* 1998) aux produits PCR a pour effet une légère séparation des profils entre *E. grandis*, *E. urophylla* et l'hybride 101 de Tm respectifs : 80°C, 79,5°C et 78°C. Cette séparation n'est cependant pas suffisante car seulement deux degrés de différence séparent les Tm de ces individus. De plus, un Tm unique est observé au sein d'un même individu hétérozygote alors que deux Tm représentatifs de chaque allèle devraient apparaître. On observe ainsi un seul Tm chez l'individu *E. grandis* hétérozygote pour le locus étudié.

Les tests réalisés sur différentes EST (06C10, 10E08, 11E10, 11F03 et 11G03) ont montré que chaque EST a sa propre température de fusion. Comme illustré sur la Figure 8, les Tm respectifs des EST 06C10, 10E08 et 11G03 sont de 80°C, 83°C et 86°C. Au sein d'une même EST, les individus étudiés ont des Tm très proches avec 1°C de variation. De plus, les individus hétérozygotes pour le locus considéré présentent toujours un Tm unique ce qui empêche

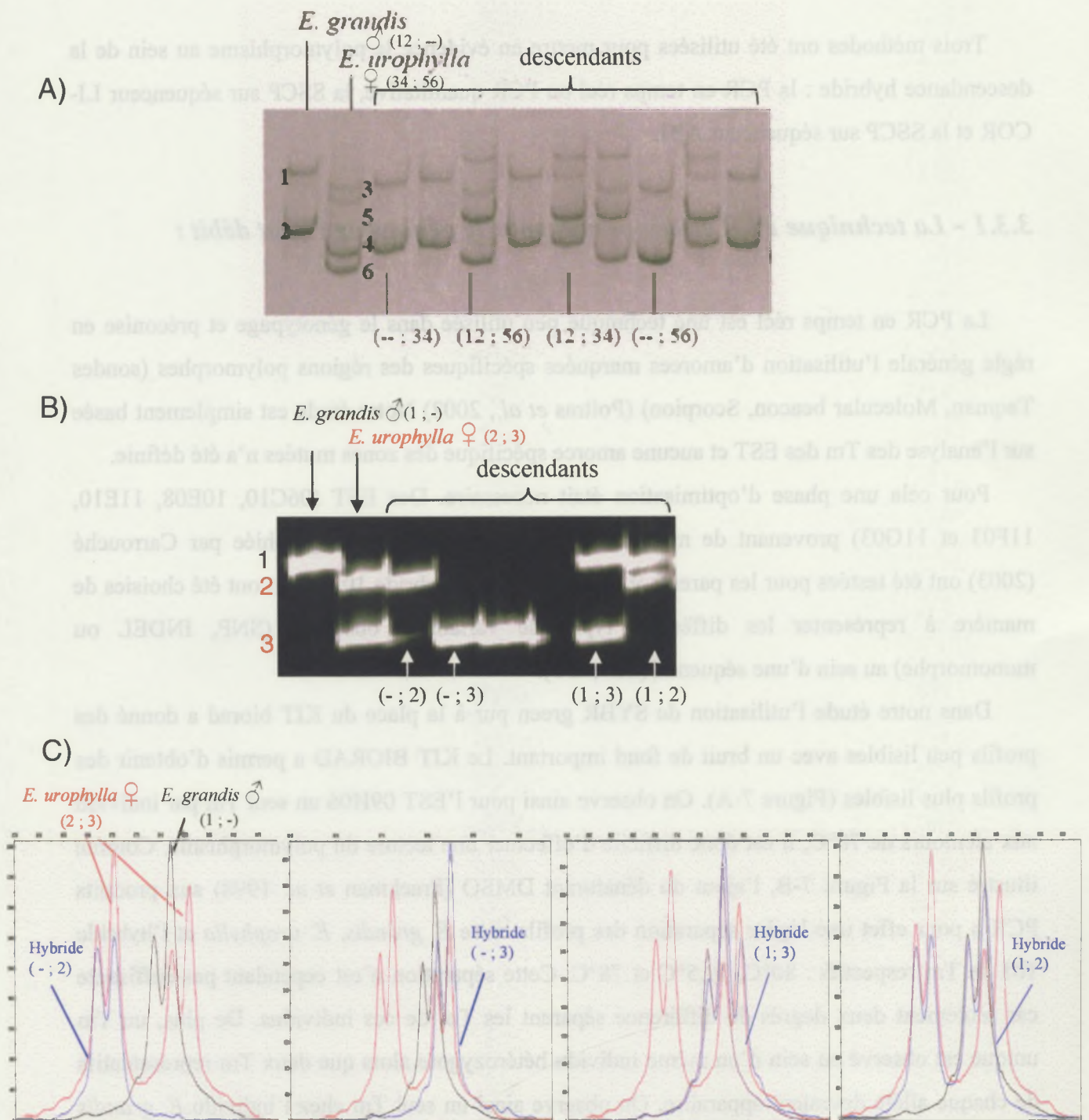


Figure 9 : Comparaison de trois profils SSCP obtenus pour l'EST 06D01, en SSCP classique (A), SSCP sur LICOR (B) et SSCP sur ABI (C).



l'identification de l'hétérozygotie pour les locus étudiés et par conséquent le génotypage des descendants.

Ainsi, lors de cette étude, les tests de PCR en temps réel en utilisant le SYBR Green n'ont pas été concluants pour révéler des variations nucléotidiques au sein d'une EST.

### **3.3.2 - Génotypage par SSCP sur LI-COR et SSCP sur ABI :**

#### **3.3.2.1 - Premiers tests effectués pour la SSCP LI-COR et ABI :**

Les premiers tests SSCP effectués sur LI-COR et ABI, ont été réalisés sur les parents et 6 descendants pour des EST déjà cartographiées par Carrouché (2003). Les premiers paramètres fixés sur LI-COR (voltage de 1000, température de 25°C, migration de 12h30) et ABI (injection de 20 sec à 15 kV, migration de 45 min à 25°C) ont donné des résultats satisfaisants.

A titre de comparaison, la Figure 9 illustre les profils obtenus pour l'EST 06D01 par SSCP classique révélée au nitrate d'argent (Figure 9-A), SSCP sur LI-COR (Figure 9-B) et SSCP sur ABI (Figure 9-C). Le profil SSCP classique (Figure 9-A) obtenu pour l'EST 06D01 (réalisé par Carrouché, 2003) montre 2 bandes chez le père 9.21 (1 2) et 4 bandes chez la mère 14.144 (3 4 5 6). Dans la descendance, 4 classes d'individus sont observées (-- ; 34), (-- ; 56), (12 ; 34), (12 ; 56) correspondant à la ségrégation des allèles parentaux. L'observation de quatre classes d'individus signifie que les deux parents sont hétérozygotes pour l'EST étudiée. Les deux allèles d'*E urophylla* (14.144) sont représentés respectivement par les bandes 3 et 4 pour l'un et les bandes 5 et 6 pour l'autre. *E grandis* (9.21) présente un allèle visible sur gel par les bandes 1 et 2 et un allèle non détectable sur gel. Cet allèle est nommé «allèle nul» car il n'est pas visualisable sur gel. L'existence d'une mutation au niveau du site de fixation du primer pour cet allèle pourrait empêcher son amplification.

Le profil SSCP obtenu sur LI-COR (Figure 9-B) pour cette EST montre une bande chez le père 9.21 (1 ; -) et deux bandes chez la mère 14.144 (2 ; 3). Contrairement à la SSCP classique qui permet de révéler les deux ADN simple brin de l'allèle (chacun adoptant sa propre conformation secondaire), dans notre étude un seul des simples brins est marqué par le

l'identification de l'hétérozygotie pour les locus étudiés et par conséquent le génotypage des descendants.

Ainsi, lors de cette étude, les tests de PCR en temps réel en utilisant le SYBR Green n'ont pas été concluants pour révéler des variations nucléotidiques au sein d'une EST.

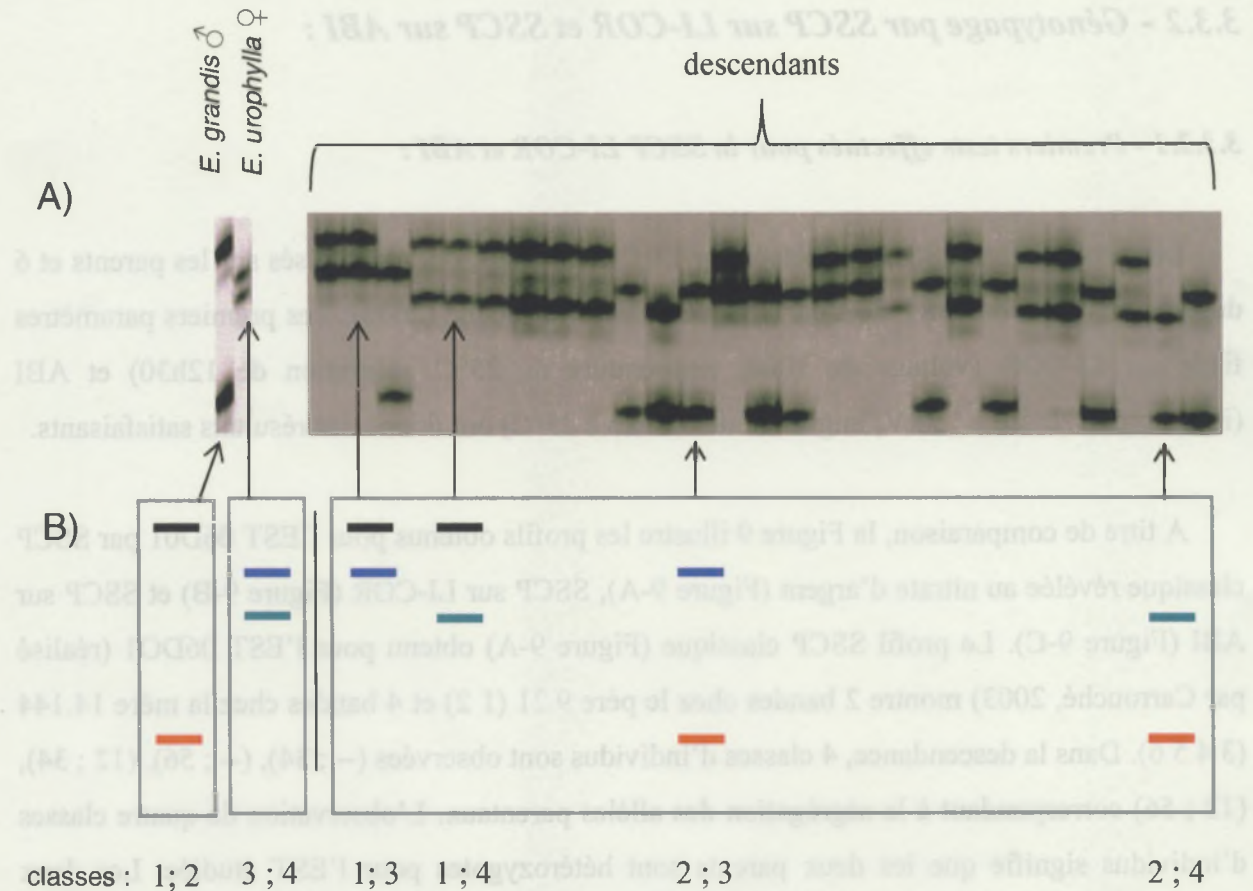


Figure 10 : Révélation du polymorphisme pour l'EST 09H06 par SSCP sur LI-COR (A) et interprétation génétique de différents profils (B).



fluorochrome ce qui explique qu'une seule bande soit visualisable pour un allèle. Les quatre classes visibles chez les hybrides testés sont donc de type (1 ; 2), (- ; 2), (1 ; 3) et (- ; 3).

Le profil SSCP sur ABI (Figure 9-C) montre un double pic chez 9.21 (noté 1 ; -) et deux double pics chez l'individu 14.144 (notés 2 et 3). Un brin est donc ici représenté par un double pic. Les allèles 2 et 3 ayant des temps de migration très proches, le profil de l'hybride 119 type (1 ; 3) apparaît sous forme d'un triple pic.

Si l'on compare les profils obtenus pour ces trois types de SSCP : les résultats sont cohérents dans l'interprétation finale des figures. Le père présente dans les trois cas un allèle nul tandis que la mère présente deux allèles visibles.

L'EST 09H06 a été passée sur LI-COR pour les deux géniteurs et 62 descendants. La lecture des profils a permis de voir que le génotypage par SSCP sur LI-COR est très résolutif (Figure 10). On identifie clairement une hétérozygotie chez les deux parents *E. grandis* de classe 1 ; 2) et *E. urophylla* de classe (3 ; 4) et quatre classes (1 ; 3), (1 ; 4), (2 ; 3) et (2 ; 4) pour les 62 descendants testés.

Pour conclure, les techniques de révélation du polymorphisme par SSCP sur LI-COR et ABI permettent de révéler efficacement du polymorphisme et donnent des résultats cohérents entre elles.

### 3.3.2.2 - Sélection des EST polymorphes par SSCP sur LI-COR :

Sur les 61 EST : 32 EST ont été passées sur LI-COR pour les 2 géniteurs et seulement 2 descendants. Normalement, la sélection d'EST polymorphes doit se faire avec plus de descendants afin de pouvoir discerner la présence éventuelle d'allèles nuls. Cependant, ce premier tri grossier a permis d'identifier un certain nombre d'EST polymorphes chez au moins un des parents.

Une deuxième étape a consisté à :

- comparer la qualité des profils obtenus pour les deux types de SSCP (sur LI-COR et ABI)
- observer si les profils obtenus pour les deux types de SSCP concordaient systématiquement.

Pour cela 13 EST identifiées comme clairement polymorphes (06E04, 06C01, 06G01, 07E09, 07G02, 10E08, 10G01, 11C08, 11D08, 11F04, 12B02, 12F08, et 12G03) chez au moins un des

Tableau 1 : Comparaison des génotypes et polymorphisme (P pour polymorphe, M pour monomorphe et x pour données manquantes) observés par SSCP sur ABI et SSCP sur LI-COR.

EST	ABI			LICOR		
	génotypes parentaux <i>E.grandis</i> / <i>E.urophylla</i>	génotypes des hybrides	polymorphisme observé	génotypes parentaux <i>E.grandis</i> / <i>E.urophylla</i>	génotypes des hybrides	polymorphisme observé
06B06	1/1, 2/2	1/2	M	non testé		x
06C01	1/2, 3/4	1/3, 1/4, 2/4, 3/4	P	1/2, 3/4	1/3, 1/4, 2/4, 3/4	P
06C05	1/1, 2/2	1/2	M	non testé		x
06E04	1/2, 1/1	1/1, 1/2	P	1/2, 1/1	1/1, 1/2	P
06G01	x	x	x	1/-, 2/-	1/-, 1/2, 2/1, -/-	P
07D07	1/1, 2/2	1/2	M	1/1, 2/2	1/2	M
07E09	1/1, 1/3	1/1, 1/3	P	1/1, 1/3	1/1, 1/3	P
07G02	1/1, 1/2	1/1, 1/2	P	1/1, 1/2	1/1, 1/2	P
10E08	1/2, 1/1	1/1, 1/2	P	1/2, 1/1	1/1, 1/2	P
10G01	x	x	x	1/2, 1/1	1/2, 1/1	P
11C08	1/1, 1/2	1/1, 1/2	P	1/1, 1/2	1/1, 1/2	P
11D08	1/2, 1/1	1/1, 1/2	P	non testé		x
11F04	1/1, 2/3	1/1, 2/3	P	non testé		x
11G03	1/1, 1/1	1/1	M	1/1, 1/1	1/1	M
12B02	1/2, 2/3	1/2, 1/3, 2/2, 2/3	P	1/2, 2/3	1/2, 1/3, 2/2, 2/3	P
12F08	1/1, 2/3	1/2, 1/3	P	x	x	x
12G03	1/2, 2/2	1/2, 2/2	P	1/2, 2/2	1/2, 2/2	P

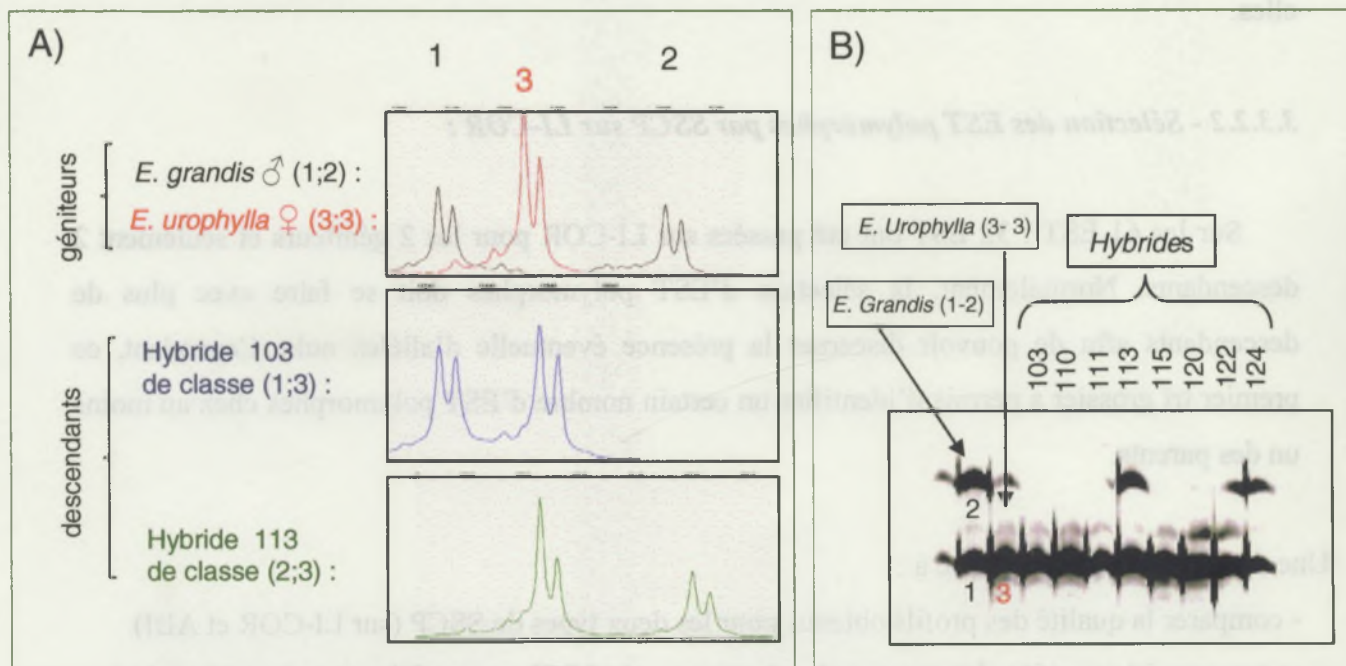


Figure 11 : Comparaison de deux profils SSCP obtenus pour la même EST 10E08, sur ABI (A) et LICOR (B).



parents et 4 EST (06B06, 06C05, 07D07 et 11G03) identifiées comme monomorphes chez les deux parents ont été passées sur séquenceur LI-COR et ABI pour les 2 géniteurs et 8 descendants.

Comme le montre le tableau 1, les génotypes observés par les deux méthodes sont identiques.

Les profils obtenus pour les 2 techniques ne sont pas toujours clairs. La SSCP LI-COR est très sensible au chargement, ainsi des bandes épaisses avec un signal intense peuvent être observées dans le cas d'un dépôt ADN trop important (supérieur à 0,3 µl). En SSCP ABI un signal saturant (pic atteignant un plateau) est observé pour un produit PCR trop concentré.

Les analyses de polymorphisme chez les parents par SSCP LI-COR ont été confirmées par les résultats de séquençage (annexe 6) : tous les résultats concordent sauf pour 5 EST 07D07, 09B10, 11E10, 11G03 et 12D06 où un polymorphisme identifié par séquençage n'a pas été révélé par SSCP sur LI-COR.

De cette analyse, 7 EST aux profils très lisibles ont été sélectionnées pour être passées sur 94 descendants en SSCP ABI (06C01, 07E09, 10E08, 11C08, 12B02 et 12F08) ou LI-COR (11D08).

### 3.3.2.3 - Génotypage des EST polymorphes sur la descendance *E. urophylla* x *E. grandis* :

L'observation des profils SSCP des descendants pour les 7 EST sélectionnées a permis l'identification de deux ou quatre classes génotypiques. Un profil de deux classes est observable quand un seul des parents est hétérozygote (Figure 11-A et B). Les profils SSCP sur LI-COR (Figure 11-B) obtenu pour l'EST 10E08 montrent deux bandes chez le père 9.21 notée (1 ; 2) et une bande chez la mère 14.144 notée (3 ; 3). Dans la descendance nous retrouvons ainsi deux classes d'individus correspondant à la ségrégation des allèles parentaux. Ainsi les individus 103, 110, 111, 115, 120, 122 appartiennent à la classe (1 ; 3) alors que les individus 113 et 124 appartiennent à la classe (2 ; 3). A titre de comparaison la Figure 11-A illustre les profils obtenus en SSCP ABI pour cette même EST. Le père 9.21 hétérozygote présente deux pics (1 et 2), alors qu'un seul pic est visible pour la mère 14.144 homozygote (3). Deux classes d'hybrides sont visualisables : tous les hybrides de type (1 ; 3) ont les 2 pics de gauche alors que les hybrides de type (2 ; 3) ont les deux pics de droite.

Tableau 2 : Interprétation des profils SSCP obtenus pour les 7 EST polymorphes et test de la ségrégation mendélienne par un test de  $\chi^2$  à 5 %. Pour  $\alpha = 5$  %, la valeur seuil du  $\chi^2$  est de 3,84 pour 1 degré de liberté (ddl) et de 7,81 pour 3 degrés de liberté.

EST	Génotypes parentaux <i>E. grandis</i> / <i>E. urophylla</i>	Génotypes des hybrides	Effectifs observés	Effectifs théoriques	$\chi^2$	ddl
06C01	(1;2) / (3;4)	(1;3) / (1;4) / (2;3) / (2;4)	49 / 34 / 36 / 25	36 / 36 / 36 / 36	2,04	3
07E09	(1;1) / (2;3)	(1;2) / (1;3)	31 / 47	39 / 39	1,64	1
10E08	(1;2) / (3;3)	(1;3) / (2;3)	41 / 48	44,5 / 44,5	0,27	1
11C08	(1;1) / (2;3)	(1;2) / (1;3)	43 / 47	45 / 45	0,08	1
11D08	(1;2) / (3;3)	(1;3) / (2;3)	39 / 45	42 / 42	0,21	1
12B02	(1;2) / (2;3)	(1;2) / (1;3) / (2;2) / (2;3)	20 / 32 / 21 / 16	22,25 / 22,25 / 22,25 / 22,25	1,58	3
12F08	(1;1) / (2;3)	(1;2) / (1;3)	40 / 54	47 / 47	1,04	1

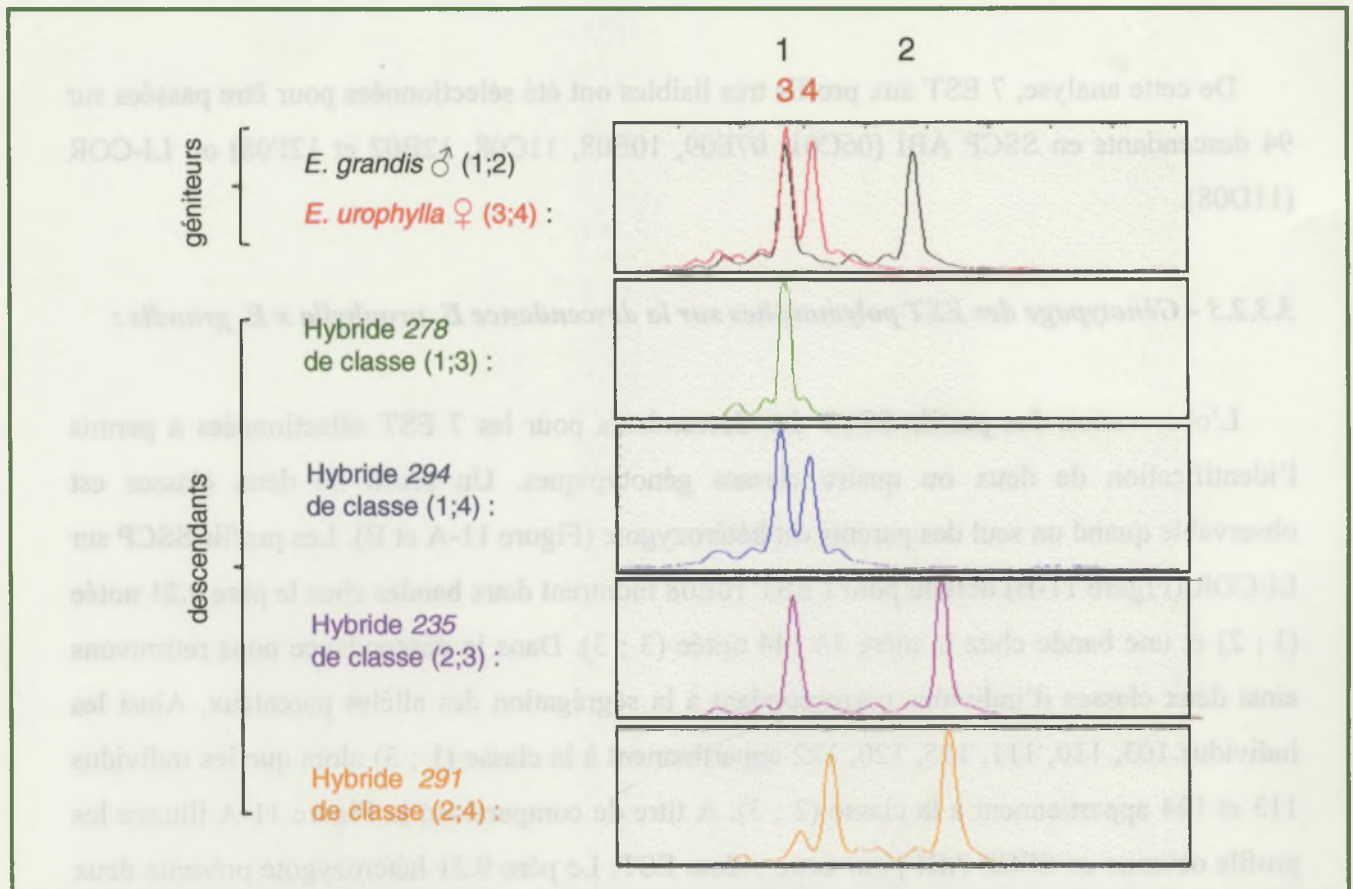


Figure 12 : Profil SSCP obtenu pour l'EST 06C01 passée sur ABI.



Lorsque les deux parents sont hétérozygotes, quatre classes sont visibles (Figure 12). Ainsi pour l'EST 06C01, la SSCP sur ABI montre deux pics chez le père (1 ; 2) et deux pics différents chez la mère (3 ; 4). Les deux parents sont donc hétérozygotes et la ségrégation allélique observée chez les descendants correspond bien à un profil à 4 classes (1 ; 3), (1 ; 4), (2 ; 3) et (2 ; 4).

Les ségrégations des allèles parentaux, observées lors du génotypage des 94 descendants du croisement *E. grandis* et *E. urophylla* pour les 7 EST, correspondent à une ségrégation mendélienne au seuil 5% (Tableau 2).

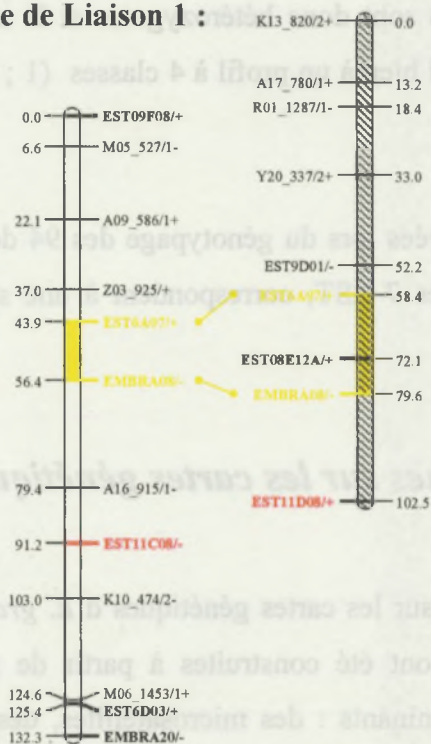
### **3.4 - Localisation des EST polymorphes sur les cartes génétiques :**

Les EST polymorphes ont été cartographiées sur les cartes génétiques d'*E. grandis* et *E. urophylla* établies par Gion (2001). Ces cartes ont été construites à partir de marqueurs dominants de type RAPD et de marqueurs codominants : des microsatellites, des gènes de fonction connues et des EST. Un total de 11 groupes de liaison sont présents chez *E. urophylla* et *E. grandis* correspondant aux 11 chromosomes du génome d'eucalyptus.

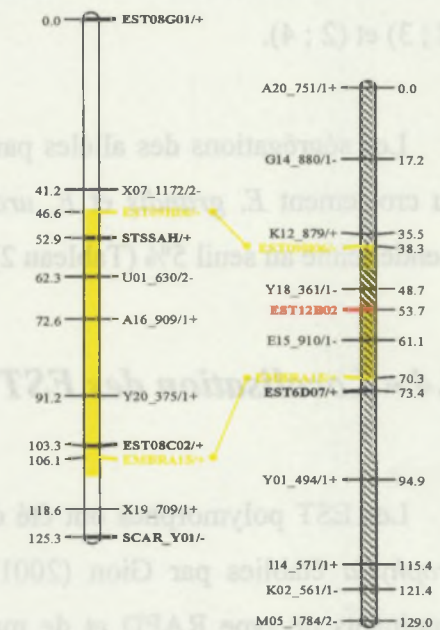
En ce qui concerne *E. urophylla*, les 5 EST polymorphes ont pu être positionnées sur les groupes de liaisons existants (Figure 13, au verso de la feuille). L'EST 11C08 a été positionné sur le GL1, les EST 07E09 et 12F08 ont été cartographiées sur le GL4, 12B02 sur le GL7 et 06C01 sur le GL10. Chez *E. grandis*, les 4 EST polymorphes ont été localisées sur les groupes de liaisons existants. L'EST 11D08 est positionné sur GL1, 12B02 sur GL3, 06C01 sur GL10 et 10E08 sur GL11. Deux EST sont polymorphes chez les deux géniteurs. L'EST 06C01 se positionne sur GL10 pour les deux géniteurs *E. grandis* et *E. urophylla*. A l'inverse l'EST 12B02 se positionne sur deux groupes de liaison différents : GL3 pour *E. grandis* et GL7 pour *E. urophylla*.

L'annexe 7 illustre le pourcentage de zones homologues entre les deux cartes *E. grandis* et *E. urophylla* après cartographie des EST étudiées.

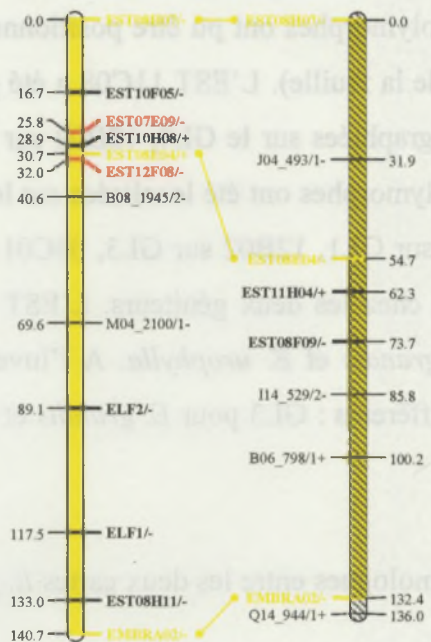
### Groupe de Liaison 1 :



### Groupe de Liaison 3 :



### Groupe de Liaison 4 :



### Groupe de Liaison 7 :

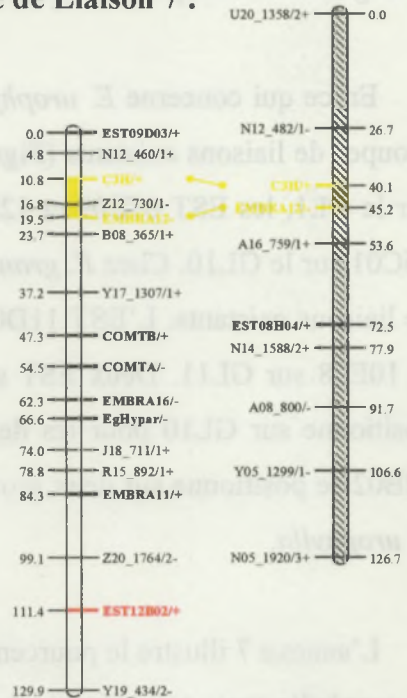
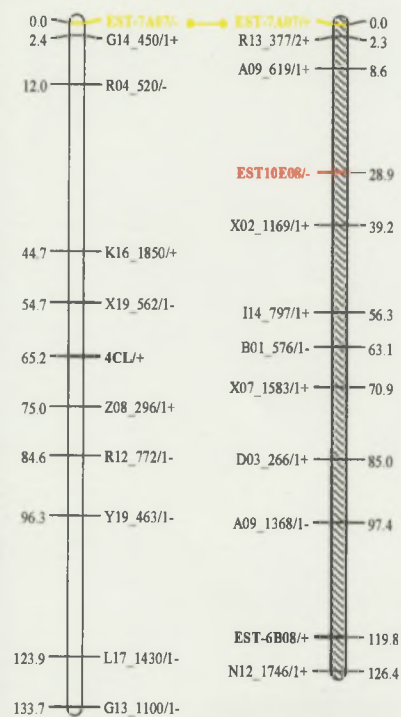
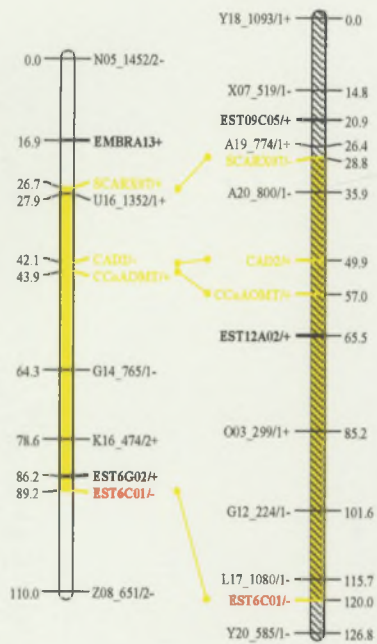


Figure 13 : Localisation des EST polymorphes sur les cartes génétiques d'*E. urophylla* et *E. grandis*.



### Groupe de Liaison 11 :



**Légende :**

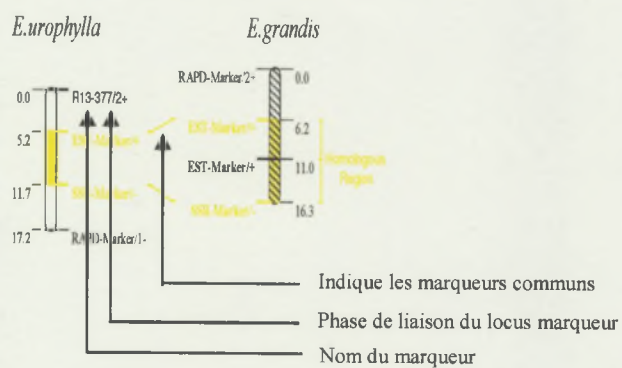


Figure 13 (suite) : Localisation des EST polymorphes sur les cartes génétiques d'*E. urophylla* et *E. grandis*.

## 4) DISCUSSION

### *La conservation des séquences entre espèces :*

Les amorces utilisées pour amplifier les EST chez *E. grandis* et *E. urophylla* ont été définies sur des séquences spécifiques d'*E. gunnii* et testées chez *E. globulus*. Ces amorces ont permis de générer 82% de produits exploitables pour la SSCP chez *E. grandis* et 74% chez *E. urophylla*. La non amplification des EST restantes peut s'expliquer par la variation nucléotidique des séquences entre différentes sections.

En effet : *E. grandis* et *E. urophylla* appartiennent à la section *Transversaria* tandis qu'*E. globulus* et *E. gunnii* appartiennent à la section *Maidenaria*. Ces résultats indiquent une meilleure conservation des séquences au sein d'une même section. Des résultats similaires avaient été obtenus par Deweer (2004) lors de son étude sur la cartographie génétique de gènes candidats pour la qualité du bois. Il avait mis en évidence une forte homologie (83%) entre les séquences d'*E. grandis* et *E. urophylla* et une plus faible homologie (68%) entre les séquences *E. grandis* / *E. gunnii* et *E. urophylla* / *E. gunnii*.

### *La PCR en temps réel :*

La PCR en temps réel basée sur l'utilisation du SYBR Green ne nécessite aucune sonde fluorescente. Le SYBR Green se fixe à n'importe quelle molécule d'ADN double brin et peut être utilisé pour détecter plusieurs produits d'amplification (Karsai *et al.*, 2001). Bien que l'on observe des  $T_m$  différents entre diverses EST, il n'est pas possible de discriminer les profils entre individus d'une même EST. Pour les géniteurs hétérozygotes pour les locus étudiés, on s'attend à observer deux  $T_m$  représentatifs de chaque allèle or un seul  $T_m$  est apparent par individu. Cette technique n'est donc pas suffisamment discriminante pour une utilisation en génotypage.

L'interprétation des résultats pour différentes EST (Figure 8), montre que la température de fusion d'un fragment n'est pas seulement dépendante de la taille de ce fragment mais que d'autres critères interviennent dans la détermination du  $T_m$ . A titre de comparaison les  $T_m$  des individus pour l'EST 06C10 est de 80°C, celui de l'EST 10E08 est de 83°C et celui de l'EST 11G03 est de 86°C; ces EST ont pourtant des tailles respectives de 200 pb, 471 pb et 182 pb.



Des études ont montré que la température de fusion dépend avant tout de la richesse en cytosines et guanines du fragment ADN considéré. En effet les trois liaisons hydrogène entre la guanine et la cytosine sont plus stables que les deux liaisons entre l'adénine et la thymine. La dénaturation d'un ADN double brin va donc être dépendante avant tout de sa composition en GC. De ce fait un SNP de type T ou A aura peu d'incidence sur le  $T_m$  par rapport à un SNP de type C ou G (Petkovski, 2006).

Pour l'étude d'un fragment, le génotypage par PCR quantitative est fortement lié à la longueur du fragment ADN considéré. En effet plus la taille du fragment ADN sera longue, plus il sera difficile de détecter des mutations par le biais du  $T_m$ , surtout pour des séquences très proches ne variant que de quelques bases. Dans notre étude, les EST utilisées dépassent les 100 pb ce qui est apparemment trop long pour obtenir des profils discriminants deux allèles.

Afin d'être discriminant dans l'analyse des  $T_m$ , Lutz *et al.* (2004) ont effectué un génotypage de différents gènes TLR (Tool like receptor) par une méthode de PCR-RFLP. Une étape de restriction est donc effectuée en amont de la PCR. La version non mutée du gène présente un site de restriction *AFIII* qui est absent dans le gène muté. De ce fait après digestion enzymatique du gène par l'enzyme *AFIII*, l'obtention de deux fragments pour le gène sauvage et d'un seul fragment pour le gène muté donne des profils  $T_m$  très discriminants (entre la version sauvage et la version mutée du gène). Sur la même idée, il serait donc intéressant pour la suite des travaux d'utiliser des enzymes de restriction, de manière à travailler sur des fragments plus courts, d'environ 20 pb. La variabilité de type SNP ou INDEL aura alors un impact beaucoup plus significatif sur la valeur du  $T_m$ .

Actuellement d'autres stratégies sont appliquées pour les techniques de génotypage par PCR quantitative. Ces méthodes sont basées sur l'utilisation de sondes fluorescentes spécifiques des zones polymorphes ciblées. L'observation d'une fluorescence va donc refléter la présence du site polymorphe ciblé. Parmi ces techniques sont principalement utilisées les méthodes Taqman (hydrolyse de sonde), Molecular Beacon (balises moléculaires) et Scorpion. La technologie Taqman (Bustin, 2000) est basée sur l'activité 5'-exonucléasique de la Taq polymérase pour hydrolyser une sonde hybridée à la séquence de l'amplicon. La sonde est constituée d'un fluorochrome et d'un quencher. L'hydrolyse de la sonde va donc permettre la libération du fluorochrome qui éloigné du quencher va pouvoir fluorescer. Les balises moléculaires (Molecular Beacons) (Tyagi et Kramer, 1996) et amorces Scorpions (Thelwell *et al.*, 2000) sont basées sur le même principe que la méthode Taqman. La sonde utilisée à une

forme d'épingle à cheveux, la portion de la sonde qui compose la boucle est complémentaire à la séquence ciblée et chacune de ces extrémités porte un fluorochrome (à l'extrémité 5') et un quencher (extrémité 3'). Pendant l'étape d'hybridation, lorsque la sonde rencontre une séquence qui lui est complémentaire, la sonde va se linéariser. Dans cette conformation, le fluorochrome est éloigné de son suppresseur, il y aura alors émission de fluorescence. La PCR en temps réel de part sa capacité à produire des résultats rapides, spécifiques et quantitatifs trouve de nombreuses applications dans différents domaines. A titre d'exemple, Steffensen *et al.* (2003) ont utilisé la PCR en temps réel pour génotyper les mutations au sein du gène humain *MBL2* (mannose-binding lectin 2) à l'aide de sondes fluorescentes spécifiques des zones mutées. Tapp *et al.* (2000) ont réalisé du génotypage à grande échelle pour le gène du récepteur d'œstrogène par les techniques Taqman et Molecular Beacon. En microbiologie, Portnoi *et al.* (2006) ont étudiés les Tm de fusion de 10 souches bactériennes différentes afin de les identifier facilement lors d'étude épidémiologiques. La PCR en temps réel est de ce fait de plus en plus utilisée dans le génotypage haut débit mais repose avant tout sur l'utilisation de sondes spécifiques des zones mutées.

### ***La SSCP sur LI-COR et ABI :***

La technique SSCP est une méthode très sensible (Bodenes, 1996) et peu coûteuse qui permet de détecter des mutations ponctuelles (Orita *et al.*, 1989). Dans notre étude, la SSCP sur LI-COR a permis de révéler du polymorphisme dans 53,3% des EST étudiées. Des résultats similaires ont été obtenus par SSCP classique chez Carrouché (2003) et Dolmare (2005) sur des EST issues de la banque de Paux (2004).

En SSCP classique, la détection de mutations ponctuelles est optimale pour des fragments de 200 à 300 pb (Markoff, 1997). Or les EST 11C08 et 10E08 identifiées comme polymorphes par SSCP sur LI-COR et ABI ont des tailles respectives de 335 et 475 pb. La taille des fragments amplifiés ne semble donc pas un obstacle pour ces deux techniques dans la détection du polymorphisme comme l'avait identifié Plomion *et al* (1999). Dans cette étude on observe que sur 7 EST contenant un intron, 5 sont identifiées comme polymorphes. Ce haut taux de polymorphisme dans les régions introniques a déjà été constaté. Champurney (2003) a montré lors d'une étude sur le gène CCR que la variabilité nucléotidique au sein du gène est en général, plus importante dans les régions non codantes par rapport aux régions codantes.



L'analyse des profils de séquençage a montré que l'analyse du polymorphisme par SSCP sur LI-COR est très pertinente. En effet cette technique permet de détecter du polymorphisme dans des séquences présentant une seule mutation ponctuelle, cas des EST 06C01, 06C10, 10G01, 11F04, 12B02 et 12G03 (annexe 6). A l'inverse pour les EST 07D07, 09B10, 11E10, 11G03 et 12D06, le polymorphisme identifié au niveau des séquences n'a pu être détecté en SSCP sur LI-COR. Cette absence de détection peut être corrélée à la méthode de marquage utilisée. En effet, pour des individus hétérozygotes, 4 bandes sont observées en SSCP classique tandis que 2 bandes sont détectées en SSCP sur LI-COR et ABI. La coloration au nitrate d'argent permet en SSCP classique de marquer les deux simples brins de l'allèle. Dans notre étude un seul brin est marqué par le fluorochrome. Si le brin marqué ne révèle pas autant de polymorphisme de conformation que son complémentaire, il est possible de n'observer qu'une seule bande chez un individu hétérozygote. Pour la suite des travaux, il serait donc intéressant de marquer les 2 amorces (Forward et Reverse) afin de comparer la qualité de polymorphisme mis en évidence.

Les deux techniques SSCP présentent néanmoins deux avantages principaux par rapport à la SSCP classique, notamment :

- la possibilité de pouvoir réaliser du multiplexage c'est-à-dire de charger deux individus dans un même puit en les marquant de fluorochromes différents (IRD700 ou 800 pour LI-COR, FAM6 ou HEX sur ABI),
- d'éviter le marquage au nitrate d'argent très coûteux en terme de temps.

#### ***Comparaison des deux techniques SSCP sur LI-COR et ABI :***

Si au niveau expérimental, la SSCP ABI est simple à réaliser (deux PCR à faire avant l'analyse sur séquenceur), la lecture des profils reste toutefois fastidieuse :

- il faut parfois réassigner le marqueur de taille afin de calibrer les profils des différents individus.
- l'analyse des résultats sur GeneMapper ne permet la comparaison que de 16 individus à la fois. Ainsi, la lecture des génotypes est relativement longue pour une descendance importante. En terme de coût le passage de la plaque sur séquenceur ABI est extrêmement onéreux (170 euros par plaque), il prend en compte le coût du marquage FAM6, l'utilisation du CAP polymère et du marqueur de taille.

La SSCP sur LI-COR à l'inverse présente une lecture très simple et globale des profils (lecture du gel) mais la manipulation est coûteuse en temps (deux PCR à réaliser, un gel à couler et un chargement ADN à effectuer).

Par rapport aux critères lecture des profils, coût et durée de la technique : la SSCP sur LI-COR semble être une technique très appropriée pour le génotypage haut débit.

### ***La cartographie d'EST :***

Lors de cette étude, l'utilisation de marqueurs codominants de type SSCP nous a permis pour l'EST 06C01 de confirmer une homologie entre les groupes de liaison 10 des deux espèces *E. grandis* et *E. urophylla*. De plus la cartographie de cette EST a considérablement augmenté la taille des zones homologues chez les cartes des deux géniteurs (la zone homologue chez *E. urophylla* et *E. grandis* à respectivement augmentée d'un facteur 30 et 12).

A l'inverse les groupes de liaison 1, 7 et 11 présentent de très faibles homologies entre les cartes d'*E. grandis* et *E. urophylla*. Il serait donc intéressant d'intégrer un grand nombre de marqueurs codominants sur ces groupes de liaison. Brondani *et al.* (2002) ont montré que 90% des microsatellites développés étaient transférables dans une même section et 80% le sont au sein du sous-genre *Symphyomyrtus*. Ces résultats révèlent une bonne conservation des génomes entre les espèces d'*Eucalyptus* appartenant au sous genre *Symphyomyrtus* ce qui permet par cartographie comparée de positionner des marqueurs orthologues sur les cartes génétiques de différentes espèces d'*Eucalyptus*. De plus, les cartes développées par Brondani *et al.* (2006) sur *E. grandis* et *E. urophylla* (basées sur l'utilisation de marqueurs codominants de type microsatellites) présentent des similitudes avec les cartes établies par Gion (2001). Ainsi, les microsatellites EMBRA 8 et EMBRA 20 du GL1, EMBRA15 du GL 3, EMBRA 2 du GL4, EMBRA 11 et 12 du GL 7 et EMBRA 13 du GL10 des cartes développées par Gion (2001) sont respectivement situés sur les groupes de liaison 6, 8, 11, 1 et 9 des cartes consensus établies par Brondani *et al* (2006).

Brondani *et al* (2006) ont eux mêmes identifié des marqueurs communs entre les cartes génétiques d'*E. urophylla* et *E. grandis* et celle d'*E. globulus* (Bundock *et al.*, 2000 ; Thamarus *et al.*, 2002). De plus, 18 gènes candidats pour l'anatomie des fibres et les caractères floraux ont été cartographiés sur les cartes d'*E. globulus* (Thamarus *et al.*, 2002). Ces gènes



Tableau 3 : Fonction et identité des EST constituant un cluster sur le groupe de liaison 4 de la carte d' *E. urophylla*. (-) signifie donnée manquante.

<b>EST</b>	<b>FONCTION</b>	<b>IDENTITE</b>
<b>07E09</b>	<b>signal transduction</b>	<b>ring zing finger</b>
<b>08E04</b>	<b>cell biogenesis</b>	<b>cellulase</b>
<b>10H08</b>	<b>transport</b>	<b>probable abc transporter</b>
<b>10F05</b>	<b>signal transduction</b>	<b>leucine rich receptor like protein kinase</b>
<b>12F08</b>	<b>-</b>	<b>-</b>

correspondant à des marqueurs codominants, peuvent servir de marqueurs « ponts » entre cartes génétiques d'espèces différentes. À terme, la localisation de microsatellites et de gènes candidats pour différentes espèces d'eucalyptus permettra donc d'effectuer une cartographie comparée au sein du sous genre *Symphyomyrtus*.

La cartographie de l'EST 12B02 reste cependant problématique car située sur deux groupes de liaisons différents : le groupe de liaison 7 chez *E. urophylla* et le groupe de liaison 3 chez *E. grandis*. L'analyse BLAST du fragment amplifié a montré que ce fragment correspond bien à la séquence de l'EST 12B02. L'EST 12B02 correspond sans doute à un gène paralogue, c'est à dire un gène qui au cours de son histoire a subi un événement de duplication. Les amorces spécifiques de l'EST 12B02 ont ainsi pu se fixer sur la copie du gène localisée soit sur le groupe de liaison 3 chez *E. grandis*, soit le groupe de liaison 7 chez *E. urophylla*. Des cas similaires ont été observés sur des locus SSR : Fisher *et al.* (1998), Peakall *et al.* (1998), Karhu *et al.* (2000) ont démontré que les mêmes amorces SSR peuvent amplifier des locus non homologues dans deux génomes différents. Il serait alors possible d'avoir des cas similaires pour des marqueurs EST. Pour la suite des travaux, afin de s'assurer que le génotypage effectué sur cette EST est fiable, la totalité de la descendance devra être génotypée.

Le positionnement des EST sur les cartes génétiques n'est qu'une étape préliminaire à l'étude de l'association entre la variabilité de gènes candidats et la variation des caractères de qualité du bois. Une étude de colocalisation EST/QTL devra donc être menée pour tester si elles expliquent tout ou partie de la variation de caractères de qualité du bois (teneur en lignines, anatomie des fibres, densité...). En effet, seules les EST dont la variabilité est liée à la variation du caractère cible pourront être utilisées en sélection assistée par marqueurs.

Ces travaux de détection de QTL pour des caractères de qualité du bois sont en cours mais n'ont pas été effectués durant ce stage. Le positionnement des EST 07E09 et 12F08 sur la carte génétique d'*E. urophylla* au niveau du groupe de liaison 4 a permis d'identifier un cluster de gènes candidats (Figure 13). Le Tableau 3 illustre les fonctions putatives de ces différentes EST. Sur les 5 EST, 2 EST (07E09, 10F05) sont impliquées dans la transduction de signal. Plusieurs QTL de densité du bois (pour différents âges) ont été positionnés par Gion (2001) sur le groupe de liaison 4 d'*E. urophylla*. Pour la suite de l'étude, il faudra préciser si la variabilité des EST est liée à la variation des caractères de densité du bois. Pour cela une analyse QTL est nécessaire en incluant au jeu de données les EST cartographiées dans notre étude. L'effet d'un



QTL et de l'intervalle de confiance sur sa position sont très dépendant de l'effectif utilisé pour l'analyse de liaison. Il sera donc nécessaire de génotyper les EST sur la totalité de la descendance pour préciser si il existe ou non une colocalisation entre ces gènes et les caractères de densité du bois.

## 5) CONCLUSION ET PERSPECTIVES

Les techniques de génotypage actuelles sont principalement basées sur l'électrophorèse de l'ADN (SSCP, DGGE, TGGE) ou l'utilisation de sondes spécifiques des zones mutées (Taqman, Molecular Beacon, Scorpion). Toutes ces techniques présentent le désavantage d'être chronophage. En ce qui concerne la SSCP, DGGE et TGGE l'investissement repose sur le coulage d'un gel et le chargement de l'ADN. Pour les autres méthodes, le design des sondes nécessite une étape préalable de séquençage des parents afin de localiser les zones variables et un design très précis ciblant la zone mutée. La PCR en temps réel basée sur la simple utilisation du SYBR Green semblait donc être une alternative très intéressante pour le génotypage haut débit. L'étude a montré que les  $T_m$  ne pouvaient varier pour des fragments trop longs. De ce fait deux techniques pourraient être développées : la première couplant la restriction de fragments et la PCR en temps réel, afin d'obtenir des fragments plus courts et de ce fait une différence de  $T_m$  plus significative. La deuxième consisterait à effectuer une SSCP en amont de la Q PCR. Ainsi, après dénaturation des produits PCR et passage immédiat dans la glace les EST variables adopteraient ainsi une conformation spécifique facilement identifiable par le  $T_m$ .

L'utilisation de la SSCP sur LI-COR et ABI pour la cartographie d'EST chez l'*Eucalyptus* sont des méthodes qui se sont avérées efficaces dans la détection du polymorphisme. Les résultats obtenus ont permis d'augmenter les zones d'homologie dans le groupe de liaison 10 et d'identifier un cluster de gènes dans le groupe de liaison 4 susceptible d'être relié à un QTL de qualité du bois. Les deux techniques ont montré une fiabilité égale dans la lecture du polymorphisme. La SSCP sur LI-COR en terme de coût et lecture du polymorphisme semble plus requise pour un génotypage haut débit. La SSCP sur ABI pourrait donc être utilisée pour la validation de résultats SSCP LI-COR.

Pour la suite de l'étude, la cartographie des EST devra se faire sur la totalité de la descendance afin de préciser la position de ces marqueurs SSCP sur les cartes génétiques. De même pour les groupes de liaison présentant de faibles homologies, le génotypage de marqueurs codominants hautement polymorphes de type microsatellites permettrait d'effectuer une cartographie comparée avec d'autres espèces d'eucalyptus afin d'identifier plus facilement des QTL de qualité du bois par l'ajout de marqueurs et de gènes cartographiés chez d'autres espèces.



## Bibliographie

**Allona I, Quinn M, Shoop E, Swope K, Cyr S, Carlis J, Riedl J, Retzel E, Campbell M.M, Sederoff R, Whetten R.W.** (1998) Analysis of xylem formation in pine by cDNA sequencing. *Proc. Natl. Acad. Sci. Plant Biology* 95: 9693-9698.

**Bodenes C, Laigret F, Kremer A.** (1996) Inheritance and molecular variations of PCR-SSCP fragments in pedunculate oak (*Quercus robur* L.). *Theor. Appl. Genet* 93: 348-354.

**Bossinger G, Leitch M.A.** (2000) Isolation of cambium-specific genes from *Eucalyptus globulus* Labill. In : Cell and molecular biology of wood formation. *BIOS Scientific, Oxford*, 203-207.

**Bouvet J.M.** (1995) Evolution de la variabilité avec l'âge et corrélation juvénile adulte dans des populations d'*Eucalyptus*. Thèse de doctorat Institut National Agronomique Paris-Grignon. 236p.

**Bradshaw H.D, Villar M, Watson B.D, Otto K.G, Stewart S, Stettler R.F.** (1994) Molecular genetics of growth and development in *Populus* III. A genetic linkage map of a hybrid poplar composed of RFLP, STS and RAPD marker. *Theor Appl Genet* 89: 167-178.

**Brondani R.P.V, Brondani C, Grattapaglia.** (2002) Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. *Mol/ Genet, Genom.* 267(3): 338-347.

**Brondani R.P.V, Brondani C, Tarchini R.D, Grattapaglia.** (1998) Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *Theor Appl Genet* 97: 816-827.

**Brondani R.P.V, Williams E.R, Brondani C and Grattapaglia D.** (2006) A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. *Plant Biology* 6: 20.

**Bundock P.C, Hayden M, Vaillancourt R.E.** (2000) Linkage maps of *Eucalyptus globulus* using RAPD and microsatellites markers. *Silvae Genetica.* 49: 223-232.

**Bustin S.A.** (2000) Absolute quantification of mRNA using real time reverse transcription polymerase chain reaction assays. *Journal of molecular endocrinology.*

**Byrne M, Marquez-Garcia M.I, Uren T, Smith D.S. and Moran G.F.** (1996) Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. *Aust. J. Bot* 44: 331-341.

**Byrne M, Murrel J.C, Owen J.V, Allen B, and Moran G.F.** (1995) An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers. *Theor. Appl. Genet* 91:869-875.

**Cabannes E.** (2004) Transférabilité d'EST impliquées dans la xylogénèse au sein du genre *Eucalyptus*. Rapport de maîtrise, Université de Montpellier II.

**Carrouché A.** (2003) Cartographie génétique de gènes candidats pour la qualité du bois. Rapport de DESS, Université de Toulouse III.

**Champurney N.** (2003) Mise en évidence de polymorphisme au sein du gène CCR chez *Eucalyptus urophylla*. Rapport de DEA, Université de Montpellier II.

**Chaperon H.** (1977) Amélioration génétique des *Eucalyptus* hybrides au Congo Brazaville. In : consultation mondiale sur l'amélioration des arbres forestiers. FAO-IUFRO, 3 *Canberra* : 1055-1069.

**Creste S, Tulmann Neto A. and Figueira A.** (2001) Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. *Plant Mol. Biol. Report* 19: 299-306.

**Devey M.E, Fiddler T.A, Liu B.H, Knapp S.J, Neale D.B.** (1994) An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. *Theor Appl Genet* 88 : 273-278.

**Diatchenko L, Lau Y-F.C, Campbell A.P, Chenchik A, Moqadam F, Huang B, Lukyanov S, Gurskaya N, Sverdlov E.D, and Siebert P.** (1996) Suppression subtractive hybridization : a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci. USA*, 93: 6025-6030.

**Dolmare.** (2005) Cartographie de genes candidats pour la qualité du bois chez l'*Eucalyptus*. Université Louis Pasteur Strasbourg.

**Eyles A, Davies N.W, Mohammed C.** (2003). Wound wood formation in *Eucalyptus globulus* and *Eucalyptus nitens*: anatomy and chemistry. *Can. J. For. Res* 33: 2331-2339.

**Fisher S.G and Lerman L.S.** (1979) Length-independant separation of DNA restriction fragments in two dimensionel gel electrophoresis. *Cell* 16: 191-200.

**Fisher S.G and Lerman L.S.** (1983) DNA fragments differing by single basepair substitution are separated in denaturing gradient gels: correspondence with melting theory. *Proc. Natl. Acad. Sci. USA* 80: 1579-1583.

**Fisher P.J, Richardson T.E, Gardner R.C.** (1998) Characteristics of single and multi copy microsatellites from *Pinus radiata*. *Theor. Appl. Genet* 96 : 969-979.

**Frackman S, Kobs G, Simpson D, and Storts D.** (1998) Betaine and DMSO : enhancing Agents for PCR. *Promega Notes Number*, 65 : p27

**Gion J-M.** (2001) Etude de l'architecture génétique des caractères complexes chez l'eucalyptus : des marqueurs anonymes aux gènes candidats. *Thèse de doctora*, Université de Rennes I.

**Gion J-M, Rech P, Grima-Pettenati J, Verhaegen D. And Plomion C.** (2000) Mapping candidate genes in *Eucalyptus* with emphasis on lignification genes. *Molecular breeding*, 6: 441-449.



**Grattapaglia D.** (2004) Found in *Eucalyptus* genetic resources will undoubtedly be a key factor to reach this goal. *Genetics and Molecular Research*, 3 (3): 369-379.

**Grattapaglia D, Bertolucci F.L.G, Sederoff R.** (1995) Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. *Theor. Appl. Genet*, 90:933-947.

**Grattapaglia D. and Sederoff R.** (1994) genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics*, 137: 1121-1137.

**Groover A, Devey M, Fiddler T, Lee J, Megraw R, Mitchell-Olds T, Sherman B, Vujcic S, Williams C, Neale D.** (1994) Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. *Genetics*, 138 : 1293-1300.

**Harvengt L, Pâques M, Philipon P.** (2000) La multiplication des pins. *Biofutur*, 199 : 22-25.

**Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalarao R, Uhlen M, Teeri T.T, Lundeberg J, Sundberg B, Nilsson P et Sandberg G.** (2001) A transcriptional roadmap to wood formation. *Proc Natl Acad. Sci USA* 98, 14732-14737.

**Joazeiro C.A and Weissman A.M.** (2000) RING finger proteins: mediators of ubiquitin ligase activity. *Cell*, 102: 549-552.

**Karku A, Dieterich J.H, Savolainen O.** (2000) Rapid expansion of microsatellite sequences in Pines. *Mol. Biol. Evol*, 17: 259-265.

**Karsai A, Müller S, Platz S and Hauser M.T.** (2001) *Biotechniques*, 32: 790-796.

**Kirst M, Carcha V, Marques C.M, Zyl L.V, Myburg A, and Sederoff R.** (2002) Individual genotypes of eucalypts show striking variation in xylem gene expression. *Abstract P548 in Plant, Animal and Microbe Genomes X Conference San Diego CA Jan*, 12-16.

**Kirst M, Muburg A.A, De Leon J.P.G, Kirst E, Scott J, Sederoff R.** (2004) Coordinated genetic regulation of growth and lignification revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of *Eucalyptus*.

**Lefort B, Rodolphe F, Charcosset A.** (1990) De nouvelles perspectives pour l'analyse génétique des caractères quantitatifs. A la recherche de locus importants. *Biofutur*, 6 : 30-7.

**LeMorvan V, Formento J.L, Milano G, Bonnet J, Robert J.** (2005) Techniques de recherche des polymorphismes génétiques. *Oncologie*, 7 : 7-16.

**Mackay I.M, Arden K.E and Nitsche A.** (2002) Real time PCR in virology.

**Maillard D.** (1978) La pollinisation contrôlée (eucalyptus, pins). Document interne du Centre Technique Forestier Tropical du Congo, 54 p.

**Markoff A, Savov A, Vladimirov V, Bogdanova N, Kremensky I and Ganev V.** (1997) Optimization of single-strand conformation polymorphism analysis in the presence of polyethylene glycol. *Clinical Chemistry*, 43:30-33.

**Marques C.M, Araújo J.A, Ferreira J.G, Whetten R, O'Malley D.M, Liu B.-H, Sederoff R.R.** AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. *Theor. Appl. Genet*, 96: 727-737, 1998.

**Marques C, Araujo J.A. et Ferreira J.G.** (1998) AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. *Theor Appl Genet*, 96 : 726-737.

**Marques C.M, Brondani R.P.V, Grattapaglia D and Sederoff R.** (2002) Conservation and syntenicity of SSR loci and QTLs for vegetative propagation in four *Eucalyptus* species. *Theor Appl Genet*, 105(2-3) : 474-478.

**Martin B, Quillet G.** (1974) Bouturage des arbres forestiers au Congo : résultats des essais effectués à Pointe-Noire de 1969 à 1973. *Bois et Forêts des tropiques* 154 : 41-57, 155 : 15-33, 156 : 61, 157 : 21-40.

**Miranda I, Pereira H.** (2002) Variation of pulpwood quality with provenances and site in *Eucalyptus globulus*. *Ann. For. Sci*, 59: 283-291.

**Myburg A.A, Griffin A.R, Sederoff R.R, and Whetten R.** (2003) Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. *Theor. Appl. Genet*, 107: 1028-1042.

**Myers R.M, Maniatis T and Lerman LS.** (1987) Detection and localization of single base changes by denaturing gradient gel electrophoresis. *Methods Enzymol*, 155 : 501-527.

**Nguyen N, Chaste B.** (2006) *Eucalyptus* et environnement. N°1-2006 Fiche n° 725.

**Orita M, Iwahana H, Kanazawa H, Hayashi K. And Sekiya T.** (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single strand conformation polymorphisms. *Proc. Natl. Acad. Sci. USA*, 86: 2766-2770.

**Paux E, Tamasloukht M.B, Ladouce N, Sivadon P, and Grima-Pettenati J.** (2004) Identification of genes preferentially expressed during wood formation in *Eucalyptus*. *Plant Mol Biol*.

**Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A.** (1998) Cross-species amplification of soybean simple sequence repeats (SSRs) within the genus and other legume genera : implications for the transferability of SSRs in plants. *Mol Biol Evol*, 15 : 1275-1287.

**Petkovski E.** (2006) Polymorphisme ponctuels de sequence et identification génétique. Etude par spectrométrie de masse MALDI-TOF.

**Plomion C, Hurme P, Frigerio J-M, Ridolfi M, Pot D, Pionneau C, Avila C, Gallardo F, David H, Neutelings G, Campbell M, Canovas F.M, Savolainen O, Bodenes C. And Kremer A.** (1999) Developing SSCP markers in two *Pinus* species. *Molecular breeding*, 5: 21-31.



**Plomion C, Leprovost G, and Stokes A.** (2001) Wood formation in trees. *Plant Physiology*, 127: 1513-1523.

**Poitras E, Houde A.** (2002) La PCR en temps réel : principes et applications. *Biology and Biotechnologie*, vol 2, N°2, December 2002.

**Portnoi D, Sertour N, Ferquel E, Garnier M, Baranton G and Postic D.** (2006) FEMS *Microbiol letter*, 259: 35-40.

**Raymond C.A.** (2002) Genetics of *Eucalyptus* wood properties. *Ann. For. Sci*, **59** : 525-531.

**Riesner D, Henco K and Steger G.** (1991) Temperature-gradient gel electrophoresis : a method for the analysis of conformational transitions and mutations in nucleic acids and proteins. In : *Advances in Electrophoresis*, 4 : 169-250.

**Rosenbaum V and Riesner D.** (1987) Temperature-gradient gel electrophoresis; thermodynamic analysis of nucleic acids and proteins in purified form and in cellular extracts. *Biophys. Chem*, 26: 235-246.

**Rychlik W, Spencer W.J, and Rhoads R.E.** (1990) Optimization of the annealing temperature for DNA amplification in vitro. *Nucl. Acids Research*, 18:6409-6412.

**Saghai-Marooif M.A, Soliman K.M, Jorgensen R.A and Allard R.W.** (1984) Ribosomal DNA spacer length polymorphism in barley : Mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci. USA*, 81: 8014-8018.

**Schuelke M.** (2000) An economic method for the fluorescent labelling of PCR fragments. A poor man's approach to genotyping for research and high throughput diagnostics. *Nature Biotechnology* VOL 18, February 2000.

**Steffesen R, Hoffmann K, Varming K.** (2002) Rapid genotyping of MBL2 gene mutations using real-time PCR with fluorescent hybridisation probes. *Journal of Immunological Methods* 278: 191-199.

**Sterky F, Regan S, Rohde M, Karlsson A, Hertzberg J, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarroel R, Montagu M.V, Sandberg G, Olsson O, Teeri T.T, Boerjan W, Gustafsson P, Ulhén M, Sundberg and Lundeberg J** (1998) Gene discovery in the wood- forming tissues of poplar : Analysis of 5692 expressed sequence tags. *Proc. Natl. Acad. Sci. Plant Biology*, 95 : 13330-13335.

**Täpp I, Malmberg I, Rennel E, Wik M, Syvänen A.C.** (2000) Homogenous scoring of single nucleotide polymorphisms: comparaison of the 5' nuclease Taq Man assay and molecular beacon probes. *Biotechniques*, 28: 732-738.

**Thamarus K.A, Groom K, Murrell J, Byrne M. and Moran G.F.** (2002) A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre, and floral traits. *Theor. Appl. Genet*, 104: 379-387.

**Thelwell N, Millington S, Solinas A, Booth J and Brown T.** (2002) Mode of action and application of scorpion primers to mutation detection. *Nucleic Acids Research*, 28: 3752-3761.

**Tozaki T, Choi-Miura N.M, Matsuo Taniyama, Kurosawa M and Tomita M.** (2002) *BMC Medical genetics* 2002, 3: 6.

**Tyagi S, Kramer F.R.** (1996) Molecular Beacons : Probes that fluoresce upon hybridization. *Nature Biotechnology*, 14: 303-308.

**Verhaegen D. and Plomion C.** (1996) Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers. *Genome*, 39: 1051-1061.

**Verhaegen D, Plomion C, Gion J.M, Poitel M, Costa.P, Kremer A.** (1997) Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. *Theor. Appl. Gene*, 95 : 597-608.

**Vigneron P.** (1991). Création et amélioration de variétés d'hybrides d'*Eucalyptus* au Congo. In IUFRO Symposium Intensive Forestry: the role of *Eucalyptus*, Durban, South Africa.

**Vigneron P. and Bouvet J-M.** (1997) Amélioration des plantes tropicales; les eucalyptus. *Editions CIRAD*.

**Vigneron P. and Delwaulle J.C.** (1990) Sylviculture clonale des eucalyptus au Congo. Actes du XIXème congrès mondial de l'IUFRO à Montréal, Canada : 19p.

**Whitcombe D, Theaker J, Guy S.P, Brown T and Little S.** (1999) Detection of PCR products using self probing amplicons and fluorescence. *Nature Biotechnology*, 17: 804-807.

**Wimmer R, Downes G.M, Evans R, Rasmussen G, French J.** (2002) Direct effects of wood characteristics on pulp and handsheet properties of *Eucalyptus globulus*. *Holzforschung*, 56: 244-252.



# *ANNEXES*

Annexe 1 : Description des EST utilisées pour l'approche gène candidat. La fonction probable des EST a été obtenue par le logiciel BLAST N. X/F désigne le niveau d'expression de l'EST dans le xylème par rapport à la feuille chez *E. gunnii*. La taille de chaque EST est donnée en paire de base (pb). Les EST pour lesquelles la fonction putative est connue sont signalées en rouge.

EST	FONCTION	IDENTITE	X/F	Taille en pb
06A02	no information	no information	3,8	137
06A06	no information	no information	2,2	306
06B02	no information	no information	2,3	244
06B03	no information	no information	2,5	241
06B06	no information	no information	1,5	232
06C01	no information	no information	1,5	327
06C05	no information	no information	2,3	151
06C06	no information	no information	1,5	198
06C09	no information	no information	3,6	343
06C10	Transport : ion transporters	putative cation transporter	3,8	197
06C12	no information	no information	2,3	340
06E02	Stress : detoxification	BAG-family molecular chaperone regulator-1	3,3	653
06E04	Metabolism : lipid metabolism	beta-ketoacyl-ACP synthase	1,1	636
06E09	no information	no information	0,9	500
06F01	no information	no information	1,7	363
06F02	no information	no information	1,3	276
06G01	Signal transduction : kinase	CBL-interacting protein kinase 16	1,8	190
06H04	Unclear classification	Photoperiod responsive protein	3,6	275
07A03	Stress : stress response	ERD4 protein	0,8	678
07B07	Transport : vacuolar transport	vacuole associated annexin	4,2	446
07D02	Transport : vesicular transport	beta-adaptin-like protein	3,1	236
07D07	no information	no information	2,3	463
07E09	Signal transduction : other	C3HC4-type zinc finger	4,7	434
07F08	no information	no information	1,3	348
07G02	no information	no information	1,3	290
08A04	no information	no information	7,3	325
08C03	Energy : electron transport	ubiquinol-cytochrome-c reductase	1,3	435
08C06	no information	no information	1,4	337
08F11	Transcription : mRNA transcription	putative bZip transcription factor PERIANTHIA	4,7	363
09B10	no information	no information	2,3	314
09C04	Stress : stress response	heat shock protein	6,5	306
09E09	no information	no information	2,4	345
10A02	no information	no information	3,2	498
10C09	no information	no information	2,1	384
10E08	no information	no information	4,5	551
10G01	no information	no information	2,2	300
11A02	no information	no information	3,5	252
11C08	no information	no information	5,8	572
11D08	Metabolism : metabolism	putative glucosyltransferase	3,2	442
11E09	Energy : Tricarboxylic acid pathway	NADP dependant malic Enzyme	12,8	419
11E10	Unclear classification	Ser/pro rich protein	3,1	291
11F02	Cell Biogenesis : cell wall	UDP-glucose:sterol 3-O-glucosyltransferase	3,4	597
11F03	no information	no information	3,8	312
11F04	Metabolism : nucleotide metabolism	2-nitropropane dioxygenase	5,2	464
11F10	no information	no information	1,7	598
11G02	Energy : glycolysis	G3PDH	3,4	421
11G03	Signal transduction : kinase	protein kinase	2,3	231
11G12	Unclear classification	chloroplast nucleoid DNA-binding protein	3,2	420
11H09	Cell Biogenesis : cell wall	laccase (diphenol oxidase)	6,2	399
12E02	Cell Biogenesis : cell wall	glycine rich protein	3,8	415
12E10	no information	no information	2,1	228
12C02	Cell Biogenesis : cell wall	arabinogalactan protein	7,2	217
12D06	Metabolism : lipid metabolism	acetyl-CoA carboxylase beta subunit	2,8	245
12D12	Cell biogenesis : cytoskeleton	myosin II heavy chain	1,7	769
12E11	no information	no information	3,2	238
12F04	Transport : other transport facilitators	NADP dependant malic enzyme	5	481
12F08	no information	no information	3,8	257
12F09	Signal transduction : GTP binding protein	rac GTPase activating protein	3,9	500
12G03	no information	no information	1,9	436
12G04	no information	no information	20	443
12G12	no information	no information	7,4	359



Annexe 2 : Tailles observées et attendues pour les 61 EST étudiées et description des séquences des amorces spécifiques utilisées. Les tailles observées signalées en rouge sont supérieures aux tailles attendues et indique la présence d'introns.

EST	Primer Forward	Primer Reverse	Taille attendue	Taille observée		T°C Hyb
				E. grandis	E. urophylla	
06A02	CGACGCAACATAACAACC	TGGTAATCAGGTAAAAGACGA	110	100	-	58
06A06	GGTTGATAAGAAGGCTGGTGA	TCTGTAGGCAGAAGATTGGA	242	250	250	62
06B02	CGGCAGCAATTGATTTTTAG	GCTGCTTGGACCTATGAGAA	120	-	-	62
06B03	CCAGCCCTGAAGAAGCTAAA	GCCAGGAGAAGATTGGATGA	141	-	150	58
06B06	TTAACTATTGAGCCCAAGACA	ATAGCCAAAATCACCAGAGG	136	150	150	58
06C01	GCTTGATAGCTCCAAAATAAGG	GTCTTTGCTTCACCACGAAC	203	250	250	58
06C05	CATAAAGAATGATCGCGCAGT	TTCACCCCTGCTGTGTGGTC	100	100	100	58
06C06	CCCAAGTCCCGTTTTACTTT	CCCAAGTCCCGTTTTACTTT	112	100	100	58
06C09	TGGAGGACAAGGTCTTTTGG	TGAGCAGTCCGCTAGTTTTCTTT	190	250	250	62
06C10	CGTCCACTCCTGACAAA	TGAGCAGTCCGCTAGTTTTCTTT	101	250	250	58
06C12	GCCGAGCTAGTGGAGAAAG	TGAGCAGTCCGCTAGTTTTCTTT	281	-	-	62
06E02	GGCTGATGGAGAAGCAAGAG	TGAGCAGTCCGCTAGTTTTCTTT	521	750	750	62
06E04	GTGGCCTTCTCAGCATTCA	TGAGCAGTCCGCTAGTTTTCTTT	151	150	150	58
06E09	CGGGCTTCTGAGAATGA	TGAGCAGTCCGCTAGTTTTCTTT	410	410	410	62
06F01	TGCTTATTTTGCAGCCATCA	TGAGCAGTCCGCTAGTTTTCTTT	191	-	-	58
06F02	GCGGATGCACTCAAAAAGT	TGAGCAGTCCGCTAGTTTTCTTT	156	-	-	58
06G01	GGGGACACAGTTGGAATACG	TGAGCAGTCCGCTAGTTTTCTTT	133	135	135	60
06H04	GAGGTGGTGGCAGGAGTG	TGAGCAGTCCGCTAGTTTTCTTT	163	175	175	60
07A03	TTGTCACGATTACCTCTGCT	ATGAGCCTGCCCTTCAAAAC	469	500	-	62
07B07	CGGGAAGTAATCCGACAAAC	ATGCCAGAGGAACAAGCAG	325	350	-	62
07D02	TGACCAAAAGAGTAGAAACCA	GCCTAAAGTGTGCGATAAAAA	139	-	150	58
07D07	GATGAGGGTGGTTGAGGACG	TCAAACATCAGCGAAGACAG	241	250	250	61
07E09	ATTGGGTGTGCTGCGGT	TCTCAGGTTTGGGGTAGGAA	302	310	310	58
07F08	ATGGTTACACCCTTGCTTTCTC	GCATATCGTGGTTCTTTCAGTT	227	225	225	51
07G02	AGATTGTTGGTCTTACATTGCT	AGACAAGAAAGAGTGGACTAT	214	250	250	62
08A04	CTGGACATGGATTGCAGATTAG	AAAGTGCCCAAGAGGAAAAGAT	131	-	-	64
08C03	CAGTGTAGTCCAGTTCATTATTC	ACTTTGGTGGTTGGTTTTG	364	375	-	62
08C06	ATCAAAACAAGCAGGGAATA	CGGCTGACTCTGGGAGGTGA	205	250	-	48
08F11	TATTGCGATGGGAAAGTTGG	CGGAGACGTGAGATGTAGTCA	153	375	375	62
09B10	CCCACCATCTGAGTGACCAA	AGAGGTTTACGCTCCCTGTT	197	200	200	59
09C04	GCGATGGATCAGGTGAAG	ATCGCCATATCAGAGACCA	219	250	250	62
09H09	GTGCTAGTTTTGCTGTGAGTT	AAAGTTCGCAATCTTTCAGCT	243	250	-	50
10A02	ACGCAGGTTTATAGGCATCTT	TTGGCTTGACTTATCCTTCTC	310	310	310	57
10C09	TTACTCAAGAAACCAACAAAG	GCTACCCATAGTCAACTCACA	251	250	250	60
10E08	GCTGGAACCTACATAAAGACCA	TGCTCATAGACGAAGACAGG	471	475	475	58
10G01	TGCATCGGTTACTCTTTAGTGC	TTTTTCGGACACAACCTAGGCC	193	200	200	60
11A02	GCTGTGCGCTTTGATGTT	TTTTTGGAGAAGAGACAAGTCA	214	-	-	58
11C08	AGGCTAAAGTGGTGGGGTG	AAGTGGAAAGGAGAAGAGG	334	335	335	59
11D08	CGGATGCTCGTGAGGATTTC	ACGGGTTTCTGAGCGGGGAC	209	250	250	62
11E09	ATCTACCCACCGTTCTCCAA	AGCAGACGCGCGATCACTCT	225	310	310	66
11E10	AGCGATGGGTGTACAAGAAGAGG	TAATCAGAAATCAGGATCAATGAA	156	150	150	58
11F02	TTACCCGCTCTACCAACTCC	CAGCGTGTCCCAACAACCATT	223	310	310	64
11F03	CGGGCAGACTTAACACGA	CTCAGCAGCGAATTAGGC	165	200	200	58
11F04	TTGAGCCAAAGCTCTGGAGTA	ACGGTGGTTAAGGAATAAGGATA	223	310	310	62
11F10	GGGAAGAGCCCAATGTAGCAC	GAAGCCATAGAAATCAAATCAC	352	500	500	63
11G02	AGTTGGCTTATGATGAGTG	GAACCGCATAGTAATTGTC	206	200	200	58
11G03	ACTTCCAGAGCCGCACCTT	CGCCGTAGCCTGTCTTACC	182	200	200	58
11G12	GTGGGATGCTTGTATGCTGATG	CCTAGTGTGGTTGAGGCTGTC	181	200	200	63
11H09	ACCCGAAATATCAAAGCATA	CGAAGACACGGAACAAGGAG	244	375	375	57
12B02	AAGCAGAGGGACAAAGC	CCAGATTTCCCGGTTCT	300	200	200	59
12B10	AGAACGGATTGTGGGTTTGG	CCGGTCCCTTCGCTAGATACA	115	-	-	58
12C02	CCACCTGACTTTGACTACATC	CGAATATGCTCATTGGAAGC	142	150	150	61
12D06	AGGAATAAGAGTGTCACTGAGG	AAATAATTTGACTTAAGGCTCC	168	-	-	61
12D12	CTGTTGGTCGATTTTCAAGTTCT	GACTTGGTCTCGGGCTTTCTAT	260	375	375	68
12E11	AACAGGGGCGGCCAAGATGCT	TGTCCTGTTCTGTGGAAATCC	153	500	500	59
12F04	AATCCACGGACTCCTGAACGC	CATCTCCAGCCCTTTGAACAC	385	500	-	64
12F08	GAAATCGAAGCTCGTTGCTG	CAGAGTGCAATGGGGAAAT	110	100	100	58
12F09	GGACCATCCTGCAATGATCTCC	CTCCAGCCTCCAACATAGCCAC	227	225	225	68
12G03	AGAAGGTCCAGGTTGGTAATG	TGGAATCAGAAGGCAGTTTA	187	235	235	58
12G04	TCCAGTTGTTGGGAGCAAGTAT	GAAGAACACCCAGCAGGTAATG	326	500	-	57
12G12	GCGATCAATGATCTCCCAATC	TTCCAACCGAAGTGACCAAGC	273	-	300	68

### Annexe 3 : **PROTOCOLE de SEQUENCAGE**

- Après PCR, déposer 2 µl sur un gel d'agarose à 2% pour vérifier la présence d'un produit de PCR et vérifier si la taille est celle attendue.
- Si le résultat est positif, ajouter de l'eau milliQ dans chaque échantillon pour amener le volume à 100 µl final.
- Installer la plaque de filtration sur l'appareil d'aspiration.
- Déposer à la pipette multi canaux, la totalité des 100 µl dans chacun des puits.
- Mettre la pompe en route, fermer le robinet d'arrivée à fond.
- Lorsque la majorité des puits sont secs, arrêter la pompe et ouvrir le robinet d'air.
- Mettre sous agitation pendant 10 min maximum.
- Récupérer la totalité des échantillons à la pipette multicanaux dans une plaque neuve type Elisa.

#### **Dosage produits PCR :**

- Doser les produits PCR au spectrophotomètre NANODROP.
- Dans une plaque PCR AB 800, déposer le volume de produits PCR correspondant à 8 ng.
- Diluer la PCR si nécessaire. Ne pas excéder un volume de 6 µl.
- Faire le MIX suivant :

- Chimie Amresham Dye Terminators	1, 5 µl
- Amorce de séquence	2,5 µl (5 pmol)
- Tampon de dilution 5X (ABI)	1,5 µl
- Eau milliQ	qsp 10 µl
- Distribuer dans chacun des puits de la plaque le volume nécessaire
- Couvrir avec un film PCR
- Lancer le programme suivant :

96°C	1 min	} X 25 cycles
96°C	10 sec	
50°C	5 sec	
60°C	4 min	

#### **Précipitation de la réaction de séquence :**

- Ajout de 1 µl d'EDTA 125 mM dans chaque puit
- Ajout de 1 µl d'acétate de sodium 3 M
- Ajout de 50 µl d'éthanol 100%
- Sceller la plaque
- Incuber 10 min à température ambiante
- Centrifuger 30 min / 2999 rcf à 4°C
- A l'arrêt : immédiatement retourner la plaque sur l'évier pour enlever le surnageant
- Centrifuger en posant la plaque à l'envers sur un papier absorbant pour éliminer le reste du surnageant 1 min à 185 rcf.
- Ajouter 70 µl d'éthanol 70 %
- Centrifuger 15 min à 2999 rcf à 4 °C
- Eliminer le surnageant comme précédemment
- Refaire un deuxième lavage à l'éthanol 70 %
- Fermer la plaque et conserver au frigo (pour le lendemain) ou -20°C (quelques jours plus tard)

**Séquencage sur Megabace :** Injection 20 sec à 2 kV et effectuer un Run de 240 min à 8 kV.



Annexe 4 : Niveau de variabilité observé par séquençage (P pour polymorphe, M pour monomorphe et x pour données manquantes) au sein des 61 EST.

EST	<i>E. grandis</i>		<i>E. urophylla</i>	
	variabilité	taille obs (pb)	variabilité	taille obs (pb)
06A02	x	130	x	140
06A06	P (1 SNP)	210	P (1 SNP)	260
06B02	P (1 SNP)	100	P (INDEL)	150
06B03	M		x	120
06B06	M	80	x	70
06C01	x	140	P (1 SNP)	130
06C05	x	150	x	330
06C06	x	260	x	70
06C09	M	190	M	190
06C10	P (1 SNP)	130	P (3 SNP)	280
06C12	P (1 SNP)	240	x	320
06E02	P (5 SNP)	750	P (4 SNP)	870
06E04	x	140	x	140
06E09	P (4 SNP)	440	M	400
06F01	x	x	x	x
06F02	x	880	x	x
06G01	x	160	x	x
06H04	x	150	P (2 SNP)	220
07A03	M	510	x	x
07B07	M	720	x	x
07D02	x	220	x	110
07D07	M	220	P (1 SNP)	220
07E09	M	290	x	310
07F08	P (2 SNP)	180	P (2 SNP)	420
07G02	M	170	x	140
08A04	x	660	x	x
08C03	M	300	x	x
08C06	x	690	x	690
08F11	P (2 SNP)	270	P (4 SNP)	420
09B10	M	160	P (1 SNP)	210
09C04	P (4 SNP)	210	P (2 SNP)	230
09H09	x	1040	x	280
10A02	M	250	P (INDEL)	260
10C09	M	190	x	280
10E08	P (INDEL)	690	M	640
10G01	P (1 SNP)	130	x	220
11A02	x	x	x	x
11C08	M	310	P (4 SNP)	x
11D08	P (3 SNP)	174	x	190
11E09	P (INDEL)	265	P (1 SNP)	270
11E10	P (1 SNP)	120	P (INDEL)	200
11F02	P (1 SNP)	230	x	310
11F03	M	150	M	160
11F04	M	240	P (1 SNP)	250
11F10	M	830	M	570
11G02	M	180	x	230
11G03	P (1 SNP)	120	M	140
11G12	P (1 SNP)	160	M	150
11H09	P (INDEL)	430	x	450
12B02	P (1 SNP)	110	P (1 SNP)	120
12B10	x	x	x	x
12C02	M	100	M	110
12D06	M	166	P (1 SNP)	170
12D12	P (2 SNP)	300	P (INDEL)	440
12E11	P (INDEL)	540	P (INDEL)	450
12F04	P (1 SNP)	450	x	380
12F08	x	680	P (2 SNP)	110
12F09	M	190	P (2 SNP)	190
12G03	P (1 SNP)	140	M	140
12G04	x	1200	x	1100
12G12	x	x	P (3 SNP)	210

Annexe 5 : Différents types de variabilité observés au sein des EST sélectionnées pour les tests de génotypage par PCR quantitative.

EST	<i>E. grandis</i>	<i>E. urophylla</i>	Taille obs (pb)
06C10	P (1 SNP)	P (3 SNP)	200
09H06	P (INDEL)	M	278
10E08	P (INDEL)	M	471
11E10	P (1 SNP)	P (INDEL)	156
11F03	M	M	200
11G03	P (1 SNP)	M	182



Annexe 6 : Comparaison du polymorphisme identifié ( P pour polymorphe, M pour monomorphe et x pour données manquantes) par SSCP LI-COR et séquençage.

EST	LICOR		séquençage	
	<i>E. grandis</i>	<i>E. urophylla</i>	<i>E. grandis</i>	<i>E. urophylla</i>
06A02	P	M	x	x
06B06	M	M	M	x
06C01	P	P	x	P (1 SNP)
06C05	M	M	x	x
06C06	M	M	x	x
06C10	P	P	P (1 SNP)	P (3 SNP)
06E04	P	M	x	x
06G01	P	P	x	x
07A03	M	M	M	x
07B07	M	M	M	x
07D02	P	M	x	x
07D07	M	M	M	P (1 SNP)
07E09	M	P	M	x
07G02	M	P	M	x
08C03	x	x	M	x
09B10	M	M	M	P (1 SNP)
10C09	M	M	M	x
10E08	P	M	P (INDEL)	M
10G01	P	M	P (1 SNP)	x
11C08	M	P	M	P (4SNP)
11D08	P	M	P (3 SNP)	x
11E10	M	M	P (1 SNP)	P (INDEL)
11F03	M	M	M	M
11F04	M	P	M	P (1 SNP)
11G02	M	M	M	x
11G03	M	M	P (1 SNP)	M
12B02	P	P	P (1 SNP)	P (1 SNP)
12C02	M	M	M	M
12D06	M	M	M	P (1 SNP)
12E11	x	x	P (INDEL)	P (INDEL)
12F08	M	P	x	P (2 SNP)
12G03	P	M	P (1 SNP)	M

Annexe 7: Taille des zones homologues au sein des deux cartes génétiques d'*E. grandis* et *E. urophylla*.

groupe de liaison (GL)	zones homologues (%)	
	<i>E. grandis</i>	<i>E. urophylla</i>
GL1	20,7	9,4
GL3	24,8	47,5
GL4	97,4	100
GL7	4	6,7
GL10	71,9	56,8
GL11	0	0



## RESUME :

Dans le cadre du programme d'amélioration génétique des eucalyptus au Congo, des recherches en biotechnologie ont été engagées pour détecter des locus impliqués dans la variation de caractères quantitatifs (QTL : Quantitative Trait Loci) pour la qualité du bois. Les travaux de génomique fonctionnelle menés par le CIRAD en collaboration avec des organismes publics (CNRS) et privé (RAIZ et ENCE), ont permis de développer une banque d'EST (Expressed sequence tag) surexprimées dans le xylème. Ces EST pourraient constituer de bons candidats pour une sélection assistée par marqueurs de la qualité du bois. Afin d'ajouter un attribut « positionnel » à ces gènes candidats « expressionnels », les chercheurs du CIRAD souhaitent étudier les colocalisations entre les QTL de qualité du bois et ces gènes sur les cartes génétiques d'eucalyptus. Dans ce contexte, l'objectif de ce stage était de tester différentes méthodes de révélation du polymorphisme nucléotidique afin de pouvoir cartographier ces gènes.

Trois méthodes ont donc été utilisées pour la cartographie génétique des EST : la PCR en temps réel, la SSCP sur séquenceur LI-COR et la SSCP sur séquenceur ABI. La mise en évidence du polymorphisme n'a pas été possible par la méthode de PCR en temps réel. A l'inverse, les techniques SSCP sur LI-COR et ABI se sont révélées très efficaces pour le génotypage haut débit. Ainsi, pour les 61 EST d'*E. gunnii* étudiées, 7 EST polymorphes ont été cartographiées sur les cartes génétiques d'*Eucalyptus grandis* et *E. urophylla* déjà existantes. L'ajout de ces EST a permis d'identifier un cluster de gènes colocalisant avec des QTL de densité du bois sur le groupe de liaison 4 et d'augmenter les zones homologues entre cartes au sein du groupe de liaison 10. L'ajout de marqueurs codominants de type microsatellites sur les cartes d'*E. grandis* et *E. urophylla* devrait permettre d'étudier la syntenie et la colinéarité à travers différentes espèces du genre *Eucalyptus*, ceci étant un prérequis pour la compréhension de l'évolution du génome de l'eucalyptus et pour le transfert d'information génétique d'une espèce à une autre.

## ABSTRACT :

In the framework of the eucalyptus breeding program in the Congo, biotechnological research are being developed to detect quantitative trait loci (QTL) involved in the control of wood quality variation. Functional genomics tools have been used by CIRAD in collaboration with a public Institute (CNRS) and private companies (RAIZ and ENCE), to develop a catalog of ESTs over-expressed in differentiating secondary xylem. These ESTs could constitute relevant candidate gene for marker-assisted breeding if they turn out to be associated with wood properties. In order to add a positional attribute to these expressional candidate genes, scientists from CIRAD aimed at studying the coincidence between wood property QTLs and these genes. In this context, the main objective of this trainee was to test different polymorphism recovery methods to map these genes. Three technologies were used: real time PCR, LICOR-SSCP and ABI-SSCP based methods. If the former fail to detect polymorphism, the two others made it possible to detect polymorphism. Out of the 61 *E. gunnii* studied ESTs, 7 were localized on either the *E. grandis* and *E. urophylla* maps. Together with other ESTs already mapped, we identified an interesting clusters of candidate genes that colocalize with a wood density QTL on linkage group 4. In addition, these 7 ESTs allowed to increase our confidence in assigning homologous linkage groups from both species maps. In the near future, the mapping of codominant Simple Sequence Repeats will made it possible to study the syntenity and colinearity across different species of the *Eucalyptus* genus, a prerequisite for the understanding of *Eucalyptus* genome evolution and the transfer of genetic information from one species to another.